



UNIVERSIDAD DE MÁLAGA

FACULTAD DE TURISMO

TESIS DOCTORAL

Realidad Aumentada bajo Tecnología Móvil
basada en el Contexto Aplicada a Destinos
Turísticos

José Luis Leiva Olivencia

Mayo, 2014

Directores:

Dr. Antonio Guevara Plaza

Dr. Carlos Rossi Jiménez

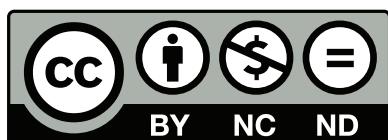


**Publicaciones y
Divulgación Científica**

UNIVERSIDAD
DE MÁLAGA

AUTOR: José Luis Leiva Olivencia

EDITA: Publicaciones y Divulgación Científica. Universidad de Málaga



Esta obra está sujeta a una licencia Creative Commons:

Reconocimiento - No comercial - SinObraDerivada (cc-by-nc-nd):

[Http://creativecommons.org/licenses/by-nc-nd/3.0/es](http://creativecommons.org/licenses/by-nc-nd/3.0/es)

Cualquier parte de esta obra se puede reproducir sin autorización
pero con el reconocimiento y atribución de los autores.

No se puede hacer uso comercial de la obra y no se puede alterar, transformar o hacer
obras derivadas.

Esta Tesis Doctoral está depositada en el Repositorio Institucional de la Universidad de
Málaga (RIUMA): riuma.uma.es

**REALIDAD AUMENTADA BAJO TECNOLOGÍA
MÓVIL BASADA EN EL CONTEXTO
APLICADA A DESTINOS TURÍSTICOS**

TESIS DOCTORAL

Presentada por

José Luis Leiva Olivencia

para optar al grado de

Doctor en Dirección y Planificación del Turismo

Dirigida por los doctores

Dr. Antonio Guevara Plaza

Dr. Carlos Rossi Jiménez



UNIVERSIDAD
DE MÁLAGA
FACULTAD DE TURISMO

Antonio Guevara Plaza y Carlos Rossi Jiménez, profesores ambos del departamento de Lenguajes y Ciencias de la Computación (Área de Lenguajes y Sistemas Informáticos) de la Universidad de Málaga.

Informan:

Que durante varios años han venido dirigiendo la Tesis Doctoral titulada “Realidad aumentada bajo tecnología móvil basada en el contexto aplicada a destinos turísticos”, realizada por el doctorando José Luis Leiva Olivencia.

Revisado el presente trabajo estimamos que puede ser presentado para su lectura y defensa al Tribunal que ha de juzgarlo. Y para que conste, autorizamos la presentación de esta Tesis en la Universidad de Málaga.

Málaga, 1 de Abril de 2014

Fdo: Antonio Guevara Plaza

Carlos Rossi Jiménez

A Isa y mis padres

Agradecimientos

Esta parte de la tesis es la última que escribo y es la que creo que más me va a costar, ya que son tantos los momentos y tantas las personas a las que agradecer su apoyo, que seguramente quedará incompleta.

En primer lugar, quiero expresar mi agradecimiento sincero a mis dos directores de tesis, Antonio Guevara y Carlos Rossi. A ambos les debo agradecer su esfuerzo, dedicación y sobre todo paciencia. Muchas gracias por haberme ayudado, aconsejado y sobre todo al tiempo que me habéis dedicado para poder llevar a buen puerto el trabajo que aquí se presenta.

También quiero desde estas líneas agradecer a mis compañeros del grupo SICUMA, en especial a Manuel Enciso por sus sugerencias, sus ideas, su pizarra en el despacho donde se discutían temas y se aclaraban conceptos. Tampoco puedo olvidar a mis compañeros José Luis y Andrés, el primero por sus constantes ánimos e interés, y al segundo por esa tranquilidad que transmite y contagia, que en definitiva hacen que se conviertan en dos grandes apoyos en mi labor docente en la universidad.

Tampoco puedo olvidar a la Facultad de Turismo y a todos los miembros que la forman. A pesar de ser un profesor asociado a tiempo parcial, me han acogido siempre como uno más en esa pequeña gran familia que forman: Iván, Rosa, Ana, Kike, Rafael,

Como es obvio, pero no menos importante también quiero agradecer tantas y tantas cosas a mi familia. En especial a Isa, unir mi vida a ella fue de lo mejor que me ha pasado en la vida. Tantas y tantas horas revisando y leyendo la memoria, corrigiéndome y ayudándome a expresar mejor lo que quería escribir. Por supuesto, agradecer a mis padres, todos sus esfuerzos y su cariño para conseguir que sus hijos pudieran estudiar como ellos no pudieron por la época que les tocó vivir. Tampoco puedo olvidar el apoyo del resto de mi familia y de la tropa trinitaria, así como de los amigos que siempre están ahí.

Tampoco puedo olvidarme de mis hijos Alejandro y Paloma, a los cuales en estas últimas semanas no les he podido dedicar el tiempo que hubiera querido, pero que me era necesario para poder terminar esta tesis.

Por último, quiero desde estas líneas hacer un homenaje a mi padre, recordando una anécdota que tuve con él unos pocos días antes de fallecer. En uno de los pocos momentos que estaba despierto en esos días, me preguntó preocupado: “Pepe Luis, ¿cómo te ha salido el examen?”, a lo que respondí tranquilizándole “Bien, Papá”. Bueno, creo que esa tesis que comencé a trabajar leyendo artículos mientras le acompañaba en la residencia donde estuvo sus últimos días, ya ha terminado, y allí donde estés, te digo “Papá, creo que ha salido bien”.

Índice

Capítulo 1. Introducción

1.1 Descripción del problema y motivación	1
1.2 Objetivos de la tesis y propuesta.....	7
1.3 Organización de la tesis.....	15

Capítulo 2. Realidad Aumentada y Turismo

2.1. Turismo y nuevas tecnologías.....	19
2.2. Estado del arte de la Realidad Aumentada Móvil	24
2.2.1. Realidad aumentada vs Realidad virtual.....	25
2.2.2. Técnicas de inmersión de los usuarios	30
2.2.3. Geolocalización	32
2.3. Componentes y aplicaciones de un sistema de Realidad Aumentada	34
2.3.1. Componentes hardware.....	34

2.3.2. Componentes software	35
2.3.3. Aplicaciones existentes de Realidad aumentada.....	36
2.3.4. Arquitecturas de sistemas de Realidad aumentada.	38
2.3.5. ULTRA.	40
2.3.6. MARCH.	41
2.3.7. ASTOR.	43
2.3.8. POLAR.	43
2.3.9. LAYAR.	45
2.3.10. LOCUS.	47
2.4. Componentes de un navegador de Realidad aumentada.....	48
2.5. Aplicaciones de la Realidad Aumentada al turismo.....	52
2.5.1. Aplicaciones y proyectos turísticos orientados al patrimonio.....	55
2.5.2. Aplicaciones turísticas orientas a guiado.....	58
2.5.3. Estado de la técnica en repositorios de contenidos.....	65
2.5.4. Plataformas de repositorios de contenidos.....	68
2.6. Análisis DAFO de la Realidad aumentada aplicada al Turismo	72
2.6.1. Debilidades	73
2.6.2. Amenazas	73
2.6.3. Fortalezas	74

2.6.4. Oportunidades	75
----------------------------	----

Capítulo 3. Sistemas de Recomendación

3.1. Introducción a los sistemas de recomendación	77
---	----

3.2. Mecanismos de Recomendación	82
--	----

3.2.1. Sistemas de recomendación colaborativos.....	92
---	----

3.2.1.1. Basado en modelos.....	95
---------------------------------	----

3.2.1.2. Basado en memoria.....	98
---------------------------------	----

3.2.2. Sistemas de recomendación basados en contenido	101
---	-----

3.2.3. Sistemas de recomendación demográficos	106
---	-----

3.2.4. Sistemas de recomendación basados en conocimiento	107
--	-----

3.3. Sistemas de recomendación aplicados al contexto.....	111
---	-----

3.3.1. Definición de contexto.....	113
------------------------------------	-----

3.3.2. El contexto en los sistemas de recomendación.....	118
--	-----

3.3.3. Paradigmas para incorporar el contexto en los sistemas de recomendación.....	124
---	-----

3.4. Técnicas de hibridación	129
------------------------------------	-----

3.5. Recomendaciones para grupos	132
--	-----

3.6. Aplicación de técnicas FCA-AFC a sistemas de recomendación	137
---	-----

3.7. Conclusiones.....	138
------------------------	-----

Capítulo 4. Organización del conocimiento

4.1. Representación del conocimiento.....	141
4.1.1. Redes semánticas.....	142
4.1.2. Representación de conocimiento basado en marcos.....	142
4.1.3. Lógicas descriptivas	143
4.2. Organización del conocimiento.....	144
4.2.1. Taxonomías	144
4.2.2. Mapas temáticos.....	145
4.2.3. Tesauros	145
4.2.4. Ontologías	147
4.3. Ontologías turísticas	149
4.3.1. Harmonise	150
4.3.2. Mondeca	151
4.3.3. Hi-Touch	151
4.3.4. QALL-ME	151
4.3.5. DERI.....	152
4.3.6. EON	152
4.3.7. cDott	152
4.3.8. Otras ontologías.....	153

4.4. Escenarios para el desarrollo de ontologías	154
4.4.1. Especificación de la ontología.....	154
4.4.2. Ciclo de vida de una ontología.....	156
4.5. Organización del perfil del usuario	157
4.5.1. Representar información del perfil del usuario	157
4.5.2. Creación del perfil inicial del usuario	160
4.5.3. Retroalimentación	161
4.5.4. Adaptación del perfil.....	162

Capítulo 5. RAMCAT. Un marco de recomendación en un sistema de realidad aumentada aplicada al Turismo

5.1. Introducción	163
5.2. Aplicación de la realidad aumentada y los sistemas de recomendación a un destino turístico	166
5.2.1. Difusión del patrimonio de un destino turístico.....	166
5.2.2. Promoción del destino.....	169
5.2.3. Guiado	170
5.2.4. Búsqueda de puntos turísticos.....	171
5.2.5. Organización del viaje.....	172
5.2.6. Funcionalidad de la propuesta	172

5.3. Arquitectura RAMCAT	175
5.3.1. Componentes de RAMCAT	179
5.3.2. Sistema de interacción o Capa de interfaz	182
5.3.3. Sistema de posición.....	184
5.3.4. Sistema de procesamiento avanzado.....	184
5.3.5. Sistema de gestión de la información	186
5.3.6. Sistema de contexto.....	187
5.3.7. Sistema de gestión de perfiles	187
5.3.8. Módulos de recomendación	188
5.3.9. Módulo gestor de recomendación	188
5.3.10. Sistema gestor de rutas.....	188
5.3.11. Sistema gestor de contenidos.....	189
5.3.12. Sistema de Trazabilidad.....	190
5.4. Ontología RAMCAT.....	190
5.4.1. Jerarquía de actividades turísticas	194
5.4.2. Diseño de la ontología.....	197
5.5. Definiciones preliminares.....	197
5.6. Representación del conocimiento de los puntos turísticos.....	205
5.6.1. Taxonomía propuesta	207

5.6.2. Atributos contextuales	214
5.6.3. Atributos no contextuales	216
5.7. Representación del conocimiento de los turistas	217
5.7.1. Creación de perfiles individuales	218
5.7.2. Creación de estereotipos.....	224
5.7.3. Adquisición del perfil de necesidades.....	226
5.8. Solicitudes de recomendación	228
5.8.1. Modos de interacción del turista	228
5.8.2. Filtrado y recomendación de puntos de interés	230
5.8.2.1. Filtrado de puntos utilizando datos contextuales	233
5.8.2.2. Filtrado de puntos basado en categorías	234
5.8.2.3. Recomendación de puntos de interés.....	235
5.9. Casos de uso del sistema	236
5.10. Evaluaciones de los turistas.....	239
5.11. Funcionamiento general del sistema	244
5.11.1. Perfil del turista.....	244
5.11.2. Generación de restricciones	246
5.11.3. Módulos de recomendación.....	248
5.11.4. Módulo de comparación	249

5.11.5. Retroalimentación de perfiles	251
Capítulo 6. Modelos de recomendación	
6.1. Introducción.....	265
6.2. Pre-filtrado contextual utilizando análisis formal de conceptos difuso (PFCBI)	267
6.2.1. Descripción	269
6.2.2. Marco teórico.....	274
6.2.3. Una lógica para el manejo de implicaciones de atributos difusos.....	278
6.2.4. Aplicación de FSL a los sistemas de recomendación contextuales	280
6.2.5. Consideraciones finales	288
6.3. Motor de recomendación basado en contenido sin memoria (SPP ^u).....	290
6.3.1. Descripción	292
6.3.2. Tipos de escalas.....	294
6.3.3. Preliminares.....	295
6.3.4. Formalización	304
6.3.5. Proceso de filtrado y recomendación.....	307
6.3.6. Ejemplo de aplicación del modelo propuesto.....	312
6.3.7. Consideraciones finales	319
6.4. Motor de recomendación basado en técnicas demográficas (SGP ^u)	320

6.4.1. Descripción	322
6.4.2. Formalización	326
6.4.3. Proceso de filtrado y recomendación	330
6.4.4. Ejemplo de aplicación del modelo propuesto	337
6.4.5. Consideraciones finales.....	344
6.5. Motor de recomendación basado en técnicas colaborativas (SRP ^u)	350
6.5.1. Descripción	351
6.5.2. Formalización	352
6.5.3. Proceso de filtrado y recomendación	359
6.5.4. Ejemplo de aplicación del modelo propuesto	362
6.5.5. Consideraciones finales.....	370
6.6. Motor de recomendación híbrido basado en historial (SHP ^u)	372
6.6.1. Descripción	373
6.6.2. Formalización	374
6.6.3. Proceso de filtrado y recomendación	378
6.6.4. Ejemplo de aplicación del modelo propuesto	380
6.6.5. Consideraciones finales.....	385
6.7. Motor de recomendación grupal	386
6.7.1. Descripción	387

6.7.2. Formalización de la representación de la información y medidas de comparación	391
6.7.3. Proceso de recomendación	397
6.7.4. Procesos de mejoras en la recomendación para grupos.....	400
6.7.5. Ejemplo de aplicación del modelo propuesto.....	402
6.7.6. Consideraciones finales	408
6.8. Creación de rutas. Trazabilidad turística	409
6.8.1. Introducción	410
6.8.2. Utilización de cadenas de Markov para la creación de itinerarios....	411
6.8.3. Probabilidades de transición de n etapas. Generación de una ruta para n visitas	415
6.8.4. Potencia enésima de la matriz transición	418
6.8.5. Trazabilidad turística y cadenas de Markov	420
6.8.5.1. Clasificación de puntos turísticos en un grafo	420
6.8.5.2. Probabilidades estables aplicadas a la trazabilidad	421
6.8.5.3. Visitas promedio de primer pasaje	425
6.8.6. Algoritmo para la creación de ruta turística	426
6.9. Consideraciones finales.....	432

Conclusiones y trabajos futuros

Conclusiones	435
--------------------	-----

Líneas de investigación y trabajo futuro	440
--	-----

Anexo I. Ontología RAMCAT 443

I.1. Taxonomía de entidades	443
-----------------------------------	-----

Anexo II. Prototipo de herramienta basada en el modelo RAMCAT

II.1. Introducción	459
--------------------------	-----

II.2. Módulos de recomendación y funcionamiento básico.....	462
---	-----

II.2.1. Registro en el sistema.....	463
-------------------------------------	-----

II.2.2. Configuración de los motores de recomendación	464
---	-----

II.2.2.1. Configuración de recomendación completa	466
---	-----

II.2.2.2. Configuración de recomendación optimizada.....	467
--	-----

II.2.2.3. Configuración de recomendación histórica.....	467
---	-----

II.2.2.4. Configuración de recomendación en cascada (Configuración avanzada)	468
--	-----

II.3. Actividades turísticas: información y evaluación.....	469
---	-----

II.3.1. Información sobre las actividades turísticas.....	469
---	-----

II.3.2. Evaluación de las actividades turísticas.....	471
---	-----

II.3.3. Características de las actividades turísticas. Expertos turísticos....	472
--	-----

II.3.4. Características de las actividades turísticas. Turistas y expectativas	475
--	-----

II.4. Recomendaciones para grupos y generación de rutas.....	478
II.4.1. Creación de grupos	478
II.4.2. Generación de rutas.....	480
II.5. Conclusiones.....	482
Bibliografía.....	485
Índice de figuras	519
Índice de tablas.....	525
Abreviaturas	529

Capítulo 1

Introducción

1.1. Descripción del problema y motivación

Internet ha provocado un importante cambio en la realización de muchas tareas y en la forma de trabajar de la sociedad. Actualmente, antes de realizar un viaje, los turistas utilizan recursos web de todo tipo para consultar servicios y actividades, que podrán realizar al llegar al destino. El volumen de información que podemos encontrar de un determinado destino turístico es habitualmente demasiado grande, por lo que se debe dedicar gran cantidad de tiempo a buscar y encontrar lo que se necesita y realmente interesa.

En la actualidad, el turismo representa un importante papel en la economía de la mayoría de los países desarrollados, estando fuertemente influenciado por las innovaciones tecnológicas (Sundbo, Orfila y

Sorensen, 2007). Las empresas y los destinos turísticos deben ofrecer nuevos instrumentos y herramientas (Buhalis y Law, 2008), además deben estar en continua adaptación a los cambios que se produzcan, ya sean por movimientos sociales, cambios en la demanda o en el comportamiento de los consumidores (Esteve y Fuentes, 2000; Castejón y Méndez, 2012).

Las oportunidades de acceso y gestión, introducidas por las tecnologías de la información y la comunicación (TIC), han acelerado la competencia entre las empresas y los destinos (Guevara et al., 2009; Juaneda y Riera, 2011). Las TIC han cambiado la forma de comercialización y difusión de la información turística de los destinos y empresas, además del modo en el que los usuarios acceden a la información, por lo que se hace necesario una inversión adecuada para conseguir aumentar la calidad de la oferta turística, mejorar los procesos y en definitiva generar ventajas competitivas (Sundbo, Orfila y Sorensen, 2007; Xiang y Pan, 2011).

Los destinos deben tener una visión amplia y dinámica, proporcionando al turista herramientas que ofrezcan información personalizada, tratando de integrar los diferentes productos turísticos existentes (Guevara, 2008; Valdes, Valle y Sustacha, 2011). A pesar de que las herramientas clásicas de información como las guías, audio-guías o páginas web pueden cubrir las necesidades básicas, el turista del siglo XXI demanda nuevas herramientas tecnológicas para obtener información con las características citadas (Guevara et al, 2009).

Actualmente, el uso de aplicaciones móviles está constituyendo una auténtica revolución en el mercado de los *smartphones* y *tablets*, suponiendo un estímulo para el desarrollo de nuevos sistemas (Lymberopoulos et al., 2011). Por tanto, las aplicaciones móviles orientadas al turismo pueden convertirse en un nuevo canal que posibilite información, promoción y comercialización turística.

Otro tipo de tecnología muy interesante es la realidad virtual, técnica que permite la generación de entornos de interacción con elementos virtuales, produciendo a los usuarios una apariencia de realidad en tiempo real (Azuma et al., 2001; Liarokapis 2006). Esta tecnología es de difícil aplicación en las empresas y destinos turísticos, porque el turista necesitaría utilizar dispositivos bastante incómodos (cascos con visores especiales, gafas, guantes, etc.).

Por otro lado, los sistemas de realidad aumentada están empezando a tomar relevancia en el turismo, ya que permiten que el mundo real, visualizado por la cámara de un dispositivo tan popularizado como un *smartphone*, se enriquezca con la inclusión de elementos virtuales, que coexisten en un mismo espacio con la imagen capturada por el dispositivo (Lashkari, Parhizkar y Mohamedali, 2010). Además, la realidad aumentada ofrece la posibilidad de mostrar información de forma intuitiva, rápida, interactiva y atractiva. Estas características hacen que su aplicación en sectores como el turismo, patrimonio, cultura y publicidad esté creciendo considerablemente. Varios estudios indican que el número de teléfonos capaces de soportar aplicaciones de realidad aumentada en el año 2015 será aproximadamente de 1600 millones

(Holden, 2009). Este dato justifica la importancia de que las organizaciones de gestión de destinos turísticos apuesten por implantar sistemas basados en esta tecnología.

Además, las herramientas de guiado e información de puntos turísticos, basados en realidad aumentada, proporcionan una nueva perspectiva de interacción que mejora la experiencia del turista en los destinos (Goh et al., 2010). Pese a la importancia que puede llegar a tener la tecnología de realidad aumentada en el ámbito turístico, el análisis de la literatura detecta carencias, que hace necesario plantear estrategias para el desarrollo de nuevos sistemas que muestren a los usuarios solo información que le resulte interesante en el momento y contexto en que lo solicita. Entre las limitaciones que presentan las actuales herramientas turísticas basadas en realidad aumentada encontramos que:

- 1) *La mayoría de las aplicaciones de realidad aumentada habitualmente muestran a los usuarios un número muy elevado de puntos de interés: esto dificulta considerablemente la selección del elemento de información deseado, especialmente teniendo en cuenta que en la mayoría de los casos el usuario está utilizando un dispositivo con una pantalla de pequeño tamaño. Este hecho hace que los beneficios potenciales de la realidad aumentada puedan verse atenuados por la cantidad de información que el usuario tiene que desechar. Se puede considerar que éste sea uno de los factores que hacen que el uso de la realidad aumentada para la búsqueda de localizaciones, actividades turísticas y ocio*

solo alcanzase en el año 2011 un 14% de los usos de este tipo de herramienta (Ashdown, 2012). Esto indica que las aplicaciones de realidad aumentada existentes no son todavía lo suficientemente usables y útiles para los turistas.

2) *No se utilizan sistemas de recomendación:* se trata de sistemas que ayudan y guían al usuario en su proceso de búsqueda de recursos o actividades. Partiendo de datos sobre diferentes actividades o ítems, el usuario proporciona sus intereses, para que el sistema, a partir de diferentes algoritmos, genere automáticamente las recomendaciones.

3) *No poseen capacidad de aprendizaje:* cuando el usuario interactúa con la aplicación, el sistema debe aprender y actualizar el perfil del turista. Para ello el sistema de recomendación deberá tener en cuenta las evaluaciones realizadas por los turistas. Dicho sistema hará que aquellas actividades que el turista evalúe positivamente sean una señal de qué tipo de recursos turísticos desea visitar. Este tipo de retroalimentación explícita presenta dificultades en aquellos casos en los que el usuario no sea proclive a valorar y evaluar las actividades. Por tanto, es necesario obtener la mayor cantidad de información sin la intervención directa de los usuarios, haciendo más eficaz el mecanismo de retroalimentación del sistema.

4) *No utilizan atributos contextuales:* los sistemas existentes no tienen en cuenta atributos contextuales como el clima, horarios, fechas, etc. Es muy importante tener en cuenta este

tipo de información para permitir descartar aquellos puntos que no cumplen criterios contextuales, consiguiendo de esta forma eliminar puntos turísticos, lo que haría más eficiente el sistema. No tiene sentido que el sistema de recomendación realice cálculos para determinar si una actividad coincide con las preferencias del turista, si no cumple los criterios contextuales. Por ejemplo, un museo que está cerrado ese día no debe aparecer como actividad recomendada, dado que haríamos perder el tiempo al turista, además de generar una desconfianza en el sistema.

5) *No tienen en cuenta las preferencias individuales para la realización de un viaje en grupo.* Los intereses y preferencias individuales son importantes a la hora de elegir qué actividad se desea realizar o visitar cuando se viaja en grupo. Las aplicaciones de realidad aumentada existentes en el mercado no permiten tener en cuenta las preferencias individuales de diferentes individuos que realizan una visita en grupo.

6) *No contemplan el grado de flexibilidad que posee cada uno de los individuos que viajan en grupo:* se trata de una dificultad muy relacionada con la anterior. En todo grupo en el que viajan diferentes individuos, encontramos personas que son más reacias a visitar puntos turísticos que no le agradan en demasía y otras que se adaptan más fácilmente.

7) *No generan rutas dinámicas a partir de un conjunto de actividades:* es muy habitual cuando se desea visitar varios puntos de interés utilizar como principal criterio la cercanía entre

ellos. A veces se deja de visitar puntos más interesantes por otros más cercanos al lugar en el que se encuentra el turista. Normalmente, las herramientas existentes tienen rutas predefinidas, pero no permiten crearlas dinámicamente.

Partiendo de estas limitaciones, enfocamos nuestra investigación a desarrollar una aproximación teórica de un sistema de realidad aumentada móvil contextual aplicada al turismo, que solucione las carencias descritas anteriormente y que pueda ser aplicada eficazmente a los destinos turísticos.

Con el desarrollo de esta tesis se establece un marco para poder implementar los sistemas de recomendación contextuales a los sistemas de realidad aumentada (los cuales están suscitando un gran interés actualmente), permitiendo a los turistas organizar de forma más eficiente su visita a un destino turístico.

1.2. Objetivos de la tesis y propuesta

El objetivo del trabajo es definir un soporte teórico para la creación y configuración de un sistema de realidad aumentada para un destino turístico, donde los usuarios puedan disponer de herramientas para planificar individualmente o en grupo visitas o rutas turísticas, teniendo en cuenta sus preferencias y contexto.

De esta forma, se plantea un método que resuelva los problemas existentes en los sistemas de realidad aumentada y su aplicación al ámbito turístico. Para ello se hace necesario adoptar técnicas que

muestren un número adecuado de puntos en el dispositivo, evitando que un exceso de información resulte incómodo al usuario, por lo que es imprescindible que la respuesta sea personalizada. Además, aunque existen técnicas de recomendación que ofrecen resultados personalizados, estos sistemas son demasiado generalistas y no tienen en cuenta diferentes parámetros contextuales que deben influir en los resultados que se recomiendan.

La utilización de técnicas de filtrado colaborativo tradicionales en dominios caracterizados por la heterogeneidad de los productos repercute negativamente en la calidad de las recomendaciones (Leiva, Guevara y Rossi, 2012). El éxito de una recomendación colaborativa dependerá principalmente de la selección de los usuarios afines a aquel que recibe las recomendaciones, ya que las preferencias de estos usuarios determinarán que un ítem sea recomendado o no. Muchas de las técnicas tradicionales se limitan a seleccionar un único grupo de usuarios con intereses similares, independientemente del tipo de producto. Esto es un punto desfavorable en el caso del turismo, ya que un sistema puede recomendar diferentes tipos de actividades como restaurantes, monumentos, jardines, etc., pero un mismo grupo de turistas puede tener intereses parecidos en restaurantes, pero gustos totalmente diferentes en otro tipo de actividades.

Dada la motivación de nuestra investigación, debemos tener en cuenta que la información que vamos a manejar es altamente subjetiva, pues contamos con información sobre opiniones, evaluaciones, percepciones, preferencias y gustos del turista. Pero también contamos

con información objetiva como son horarios, clima y personas con las que se realiza la visita, los cuales son datos que no son utilizados en la mayoría de los sistemas existentes.

El objetivo principal de la tesis expresada formalmente es:

“Ofrecer un marco teórico sobre la arquitectura de un sistema de realidad aumentada basado en el contexto aplicado al turismo, que permita implantarse en cualquier destino turístico, ofreciendo al turista una respuesta personalizada en función de sus preferencias en cada momento, permitiendo además al destino conocer los diferentes perfiles de turistas que lo visitan”.

Entre las principales contribuciones de la tesis destacamos:

- a) Analizar y evaluar detalladamente diferentes herramientas existentes de realidad aumentada y de técnicas utilizadas en los sistemas de recomendación, así como identificar las carencias existentes para su aplicación en el ámbito turístico.
- b) Ofrecer una nueva visión de los sistemas de realidad aumentada orientados al contexto del turista y a las visitas en grupo en un destino turístico.
- c) Mejorar los procesos de adquisición de información para la construcción de perfiles de usuarios y perfiles de grupos. Para ello se requiere utilizar técnicas que minimicen la interacción del usuario, mediante la aplicación de sistemas de inferencias que permitan obtener los perfiles adecuados del turista. Con este objetivo se definirán parámetros implícitos y explícitos que

permitirán analizar el comportamiento de los turistas al usar el sistema.

d) Aplicar reglas de dependencia que relacionen atributos contextuales con características de puntos turísticos. Por ejemplo, si está lloviendo, parece razonable que el sistema no debe recomendar puntos turísticos en lugares abiertos y sí en lugares cerrados. De esta forma podrán obtenerse recomendaciones más eficientes.

e) Diseñar un algoritmo que permita aplicar recomendaciones a un sistema de realidad aumentada para destinos turísticos. Los modelos de recomendación deben ser eficientes incluso en situaciones donde la información histórica sobre valoraciones o acciones anteriores del turista no existan o sea escasa. Por tanto, se debe combatir la deficiente calidad de la mayoría de las recomendaciones, cuando la información sobre las preferencias del turista sea deficiente.

f) Utilizar técnicas para la creación dinámica de rutas a partir de un conjunto de puntos turísticos recomendados, teniendo en cuenta el contexto del turista, atributos relacionados con la visita que desea realizar (tiempo, coste económico, etc.) y rutas generales del destino.

g) Aplicar mecanismos que permitan modificar de forma automática el perfil del turista teniendo en cuenta las visitas que realiza.

h) Establecer marcos que utilicen una estrategia dinámica, que permita identificar, para cada turista, el grupo al que

pertenece cuando realice una petición de recomendación de un tipo de producto.

- i) Utilizar técnicas de recomendación a grupos de usuarios que viajan juntos.
- j) Establecer un marco teórico basado en los modelos creados y desarrollar un prototipo aplicado al destino Costa del Sol.

Para lograr estos objetivos, el sistema debe recomendar los puntos de interés turísticos a visitar en función de las preferencias y características del turista, así como del contexto en el que se encuentra. Para conseguirlo se proponen diferentes algoritmos de recomendación que realizarán las recomendaciones en cada situación, teniendo en cuenta no solo los criterios o preferencias del turista, sino también siguiendo criterios y valoraciones de otros usuarios.

Para conseguir sistemas de recomendación más eficaces será necesario mejorar el proceso de retroalimentación, consiguiendo recopilar la mayor cantidad de información posible del turista de forma implícita. La mayoría de los sistemas de recomendación están basados en valoraciones explícitas, las cuales son un elemento indispensable, pero no es menos cierto que la información implícita puede complementar y aumentar la calidad de las recomendaciones. Los valores implícitos deben convertirse en valoraciones explícitas para ayudar a los sistemas de recomendación a enriquecer la experiencia y satisfacción del turista.

Con objeto de dar soporte a estos mecanismos de recomendación se plantea la arquitectura de un sistema llamado RAMCAT (Realidad Aumentada Móvil Contextual Aplicada al Turismo), que permitirá

realizar recomendaciones, a las que los usuarios podrán acceder a través de diferentes dispositivos. Este sistema permitirá recomendar elementos muy diversos como actividades, lugares, monumentos, restaurantes, etc., de acuerdo a las necesidades del usuario y al contexto en el que se encuentra.

Tanto la recomendación como la generación de rutas se realizan dinámicamente a partir de la información almacenada sobre los usuarios, grupos, recursos e historial de visitas realizadas.

RAMCAT guiará al turista recomendándole aquellas actividades que se encuentran en la zona en la que está, así como dentro de un radio indicado por él mismo, ofreciendo recomendaciones sobre qué debe visitar, en qué restaurantes puede comer y tiempos estimados, realizando un seguimiento de sus acciones mientras interactúa con el sistema.

Para facilitar la visita se ha desarrollado un método que permite configurar una ruta a partir de la selección por parte del turista de los puntos que desea visitar. Esta opción incluye la posibilidad de configurar diferentes atributos, cómo la duración estimada para una actividad, lo que facilitará que la ruta se adapte a estos condicionantes.

Cuando varios turistas viajan en grupo RAMCAT permitirá configurar quiénes son los miembros, por lo que la fase de recomendación tendrá en cuenta las preferencias de todos los componentes del grupo. Dichas recomendaciones podrán tener en cuenta

los rasgos de flexibilidad y de adaptación de los usuarios, previamente definidas.

El sistema RAMCAT se compone de dos herramientas: RAMCAT web y RAMCAT móvil. Una de las principales funciones de la aplicación web es recoger información específica del usuario sobre sus preferencias. De esta forma, se irán identificando los intereses que permitirán no solo conocer sus gustos, sino también asignarlo a un tipo de visitante o estereotipo, según diferentes características socio-demográficas. Posteriormente, el sistema asocia estas valoraciones a los diferentes conceptos y categorías definidos en el mismo. Esta información es esencial para poder personalizar adecuadamente las recomendaciones en la aplicación móvil, teniendo en cuenta las relaciones semánticas existentes entre los diferentes conceptos.

A partir de esta información, los diferentes mecanismos del sistema realizarán una recomendación basándose en las preferencias definidas. Como hemos descrito anteriormente, uno de los principales inconvenientes de los sistemas de RA existentes, es que el número de puntos que aparecen en la pantalla es muy elevado. RAMCAT móvil evitará este inconveniente mostrando en pantalla un máximo de cinco puntos con mejor valoración o predicción de interés para el turista, pudiendo éste solicitar una ruta entre estos puntos que optimice el tiempo en los desplazamientos y tenga en cuenta los horarios de apertura y cierre, etc. Por supuesto, se podrá modificar manualmente la ruta o solicitar una nueva recomendación si se considera oportuno.

RAMCAT móvil plantea varias formas básicas de uso:

- Modo aumentado: el turista visualiza en la pantalla la realidad aumentada apareciéndole los puntos que son recomendados, teniendo en cuenta el contexto, las características y preferencias del turista que usa el sistema.
- Modo lista: se visualiza una lista de puntos turísticos ordenados teniendo en cuenta las preferencias del turista. Éste podrá recorrer ésta lista, así como obtener una descripción de los puntos recomendados. Una vez que selecciona el punto que desea visitar, la imagen real se visualizará aumentada con el punto seleccionado, lo cual le servirá para guiarse, informando, en el caso de que la distancia no sea cercana, qué medios de transporte existen para llegar hasta ese punto.
- Modo ruta: es una generalización del método anterior. Se selecciona de la lista aquellos puntos que se desea visitar y el tiempo estimado que va a dedicarse a realizar la visita. El sistema le mostrará la imagen real, apareciendo los puntos numerados por orden de visita recomendada, que conformarán la ruta teniendo en cuenta las distancias entre ellos, horarios e itinerarios que realizan la mayoría de los turistas que visitan el destino.

La propuesta de la tesis es aplicar a sistemas de realidad aumentada técnicas que permitan combinar diferentes sistemas de recomendación utilizando diversas fuentes de información. Entre ellas, habrá que tener en cuenta las preferencias del turista que solicita la recomendación, su historial de visitas, su contexto y las valoraciones realizadas por

visitantes con similares gustos. El sistema debe definir estrategias para adquirir información y obtener recomendaciones de calidad.

La implantación de herramientas innovadoras en los destinos turísticos potenciará las ventajas competitivas del sector turístico. El turista del siglo XXI demanda cada vez más información adaptada a sus preferencias, por lo que las empresas y las instituciones de gestión del turismo deben proporcionar herramientas para prestar un mejor servicio y ofertar aquellos productos que son más afines al turista.

1.3. Organización de la tesis

Los capítulos de la presente tesis se han distribuido en dos secciones bien diferenciadas:

- a) Estado del arte.
- b) Aportaciones realizadas.

La sección correspondiente al estado del arte y al estudio de los antecedentes se compone de tres capítulos en los que se estudian los sistemas de realidad aumentada, los sistemas de recomendación y las ontologías relacionadas con el ámbito turístico.

El capítulo 2, denominado “Realidad aumentada”, hace un recorrido por el estado del arte relacionándolo con la propuesta que se planteará a lo largo de los posteriores capítulos. Se particulariza el estudio a los entornos y situaciones que requieren movilidad del usuario.

En el capítulo 3, “Sistemas de recomendación”, se hará una revisión en profundidad de las técnicas de recomendación existentes, su funcionamiento, ventajas e inconvenientes. Se prestará especial atención a los sistemas de recomendación colaborativos, basados en contenido, demográficos, contextuales y grupales, pues serán los modelos que utilizaremos para alcanzar los objetivos presentados.

Este bloque termina con el capítulo 4, denominado “Organización del conocimiento”. Se estudia el estado del arte sobre la representación y organización del conocimiento. Además, se enumeran diferentes ontologías relacionadas con el ámbito turístico. Para finalizar se definen mecanismos para el desarrollo de ontologías y formas para representar información turística, así como un estudio detallado sobre la representación de la información del perfil del usuario.

La segunda sección consta de dos capítulos, donde se especifican las aportaciones realizadas. Se presentará el marco teórico en el que se inscribe el sistema RAMCAT, así como las características de los diferentes motores de recomendación propuestos.

En el capítulo 5 denominado “RAMCAT. Un marco de recomendación en un sistema de realidad aumentada aplicada al turismo”, se presenta una descripción de la arquitectura y los módulos del sistema. Esta descripción abarcará tanto los componentes estáticos como los componentes dinámicos. Se definen diferentes conceptos formalmente, proporcionando un marco conceptual común.

El capítulo 6 se denomina “Modelos de recomendación” y es el núcleo fundamental de la tesis. En él se describen los diferentes módulos de recomendación que aporta la propuesta:

- a) Módulo de recomendación basado en contenido sin memoria.
- b) Módulo de recomendación basado en técnicas demográficas.
- c) Módulo de recomendación colaborativo.
- d) Módulo de recomendación basado en contenido con memoria.
- e) Módulo de recomendación híbrido basado en historial
- f) Módulo de recomendación grupal.
- g) Módulo de generación de rutas.

Para finalizar se presentan las conclusiones y trabajos futuros en los que se sintetizan las principales aportaciones realizadas y se exponen las líneas de investigación futuras a partir de este trabajo. Después de ellas se aportan dos anexos:

Anexo I: donde se incluye el glosario de entidades que componen la ontología en la que hemos basado nuestro trabajo.

Anexo II: que describe un prototipo de aplicación móvil en el que se ilustran los mecanismos de recomendación descritos en la tesis, que dará soporte a recomendaciones a turistas que visitan la Costa del Sol Occidental. Además, se describen las diferentes formas de configuración que tiene el turista para solicitar las recomendaciones.

Para finalizar se incluye un índice de figuras, tablas y las referencias bibliográficas citadas y estudiadas a lo largo del desarrollo de esta tesis doctoral.

Capítulo 2

Realidad Aumentada y Turismo

2.1. Turismo y nuevas tecnologías

Durante las últimas décadas, el turismo ha experimentado una continua expansión y diversificación, convirtiéndose en uno de los sectores económicos de mayor envergadura y crecimiento del mundo.

En el informe “*Panorama OMT del turismo Internacional*” (OMT, 2012), se afirma que la llegada de turistas internacionales en el mundo ha registrado un crecimiento prácticamente ininterrumpido, pasando de los 277 millones de 1980 a los 1000 millones actuales.

Según *Tourism Towards 2030* (UNWTO, 2011), un estudio actualizado de la OMT sobre previsiones y análisis de tendencias a largo plazo, se cree que el número de llegadas de turistas internacionales

crecerá un 3,3% al año de media entre 2010 y 2030, lo cual implica que se alcanzará un total de 1800 millones de llegadas en 2030.

La forma más completa de medir la importancia económica tanto del turismo receptor como del turismo interno en la economía es mediante la aplicación del documento “*Cuentas satélites de turismo: Recomendación sobre el marco conceptual*” (Naciones Unidas, 2008), aprobado por la Comisión de Estadística de las Naciones Unidas. Partiendo de los datos existentes, se estima como valor medio un 5% la contribución del turismo al producto interior bruto (PIB). Además, la contribución del turismo al empleo se estima entre el 6% y el 7% del total de puestos de trabajo, incluyendo tanto los puestos directos como indirectos (OMT, 2012).

La mayoría de los turistas visitan destinos situados dentro de su propia región¹, así aproximadamente un 80% de las llegadas a un destino pertenecen a turistas de la misma región. Analizando los datos ofrecidos en el informe “*Panorama OMT del turismo Internacional*” (OMT, 2012), podemos afirmar que existe un flujo muy importante de turistas de las economías desarrolladas al resto de las economías.

Los países que más gastan en turismo internacional son: Alemania, Estados Unidos, China, Reino Unido, Francia, Canadá, Rusia, Italia, Japón y Australia. A su vez, la mayoría de estos países se encuentran entre los principales consumidores en el uso de las tecnologías de la información y la comunicación (CIA, 2012). Por tanto, podemos afirmar

¹ División realizada por la OMT en base a la morfología continental y cultural. Las regiones en que la OMT divide el mapa turístico del mundo son: Américas, Europa, Oriente Medio, África, Asia Meridional y Asia Oriental/Pacífico.

que los países que gastan más en turismo coinciden con los que poseen una mayor proporción de usuarios de Internet y un mayor porcentaje de número de líneas móviles respecto a su población.

Por consiguiente, dado que Internet es fundamental para promocionar y comercializar productos turísticos y que la mayor parte de los turistas utilizan tecnologías de la información y la comunicación, se hace necesario posicionar los destinos adecuadamente, para conseguir llegar eficientemente a ese conjunto de turistas.

Centrándonos en España, desde hace muchas décadas nuestro país es un importante receptor de turismo internacional, siendo la aportación del sector turístico al PIB nacional un factor clave. Dada la relevancia que para la economía española tiene el sector turístico, así como el cambio en comportamientos, hábitos y actitudes que manifiestan los turistas en los últimos años, resulta indispensable realizar una fuerte inversión en nuevas tecnologías, para conseguir que el turismo en España sea competitivo y siga manteniéndose a la cabeza de los destinos turísticos internacionales. (IET, 2010; Exceltur, 2012; Leiva et al., 2012).

El turismo es un sector extraordinariamente sensible a la adaptación de las nuevas tecnologías (Castejón y Méndez, 2012). Su proyección beneficia la captación de nuevos procedimientos que emanan de otros campos, incentivando la aplicación de los recursos más innovadores, convirtiendo de esta forma al turismo en un sector que marca tendencias.

Las TIC siempre han estado muy relacionadas con el turismo, la creación de los CRS (*Computerised Reservation System*) y GDS (*Global Distribution System*) demuestran que es un sector muy innovador en la aplicación de las nuevas tecnologías. No obstante, la aparición de Internet ha sido la verdadera revolución tecnológica en el ámbito turístico. Mientras los CRS y los GDS han permitido desarrollar y globalizar la disponibilidad de los servicios turísticos, Internet ha permitido a los consumidores finales disponer de los productos directamente, redefiniendo el sistema comercial y la noción de canal distributivo de los servicios turísticos (Guevara et al, 2009).

Actualmente es muy habitual utilizar el término Turismo 2.0 (Sigala, 2011; Liburd, 2012) basado en la interacción del cliente final y las diferentes aplicaciones disponibles en la web. Mediante ellas se permite no sólo acceder a la información de productos y destinos, sino que el propio turista es capaz de diseñar su propio plan de viaje, de forma que pueda comprar o reservar en la red diferentes productos (hoteles, entradas para el teatro, vuelos, etc.).

El uso de Internet se ha generalizado entre la población de una forma muy importante y se ha convertido en una herramienta indispensable para el turista a la hora de organizar y elegir su viaje (Guevara, Aguayo y Gálvez, 1997; Guevara et al, 2009; Leiva, Guevara y Rossi, 2012). Y esto no es algo puntual, sino que los datos de los últimos años señalan que es una tendencia creciente. Los resultados presentados en la conferencia de comercio electrónico europeo (ECC

European E-commerce Conference) celebrada en el año 2012², muestran que una gran parte del comercio electrónico que se lleva a cabo proviene del sector turístico y de ocio. Además, la gran mayoría de los turistas que nos visitan utilizan Internet para diseñar y organizar su estancia en España (Gago et al, 2009).

Las nuevas tecnologías aplicadas al sector turístico son por tanto, absolutamente necesarias si se quiere seguir manteniendo las excelentes cifras de turistas que visitan España, para de esta forma seguir aportando riqueza a la economía del país.

Actualmente se comienza a aplicar en el sector turístico sistemas basados en realidad virtual (RV) y realidad aumentada (RA). En concreto, los sistemas de realidad virtual pueden definirse como aquellos en los que todos los elementos que percibe el usuario son sintéticos y permiten una interacción en tiempo real (Aukstakalnis y Blatner, 1992). En cambio, en los sistemas de realidad aumentada, el mundo real se enriquece con la inclusión de elementos virtuales, coexistiendo estos con el mundo real en un mismo espacio (Azuma, 2001). La realidad aumentada se ejecuta interactivamente y en tiempo real, posicionando de forma adecuada los elementos virtuales en el entorno real.

Ambas técnicas se presentan como tecnologías muy interesantes para el turismo, ya que pueden diseñar diferentes situaciones y ambientes (Leiva, Guevara y Rossi, 2012). Además, se pueden explotar las capacidades de almacenamiento que posee un sistema informático,

² <http://www.eec-conference.com/>

permitiendo guardar, si fuera necesario, información sobre actitudes y comportamientos de los usuarios.

La realidad aumentada es una herramienta que permitirá mejorar no solo la visualización de la oferta y calidad de nuestros destinos turísticos y nuestra oferta cultural, sino que también enriquecerá el tipo de experiencia turística del visitante (Leiva et al., 2012).

2.2. Estado del arte de la Realidad Aumentada Móvil

Los comienzos de la realidad aumentada podemos encontrarlos en los años sesenta, pero no fue hasta la década de los noventa, cuando pasó a considerarse como una línea de investigación importante. Es una de las ramas tecnológicas más innovadoras, que actualmente se está empezando a utilizar en ámbitos tan diferentes como la arquitectura, educación, psicología, publicidad o ingeniería (Liarokapis, 2006). El turismo no debe ser ajeno a este fenómeno y son muchos los proyectos vinculados al patrimonio, transporte, entre otros muchos sectores, que consideran que este tipo de tecnología puede convertirse en un importante elemento para elevar los beneficios y obtener ventajas competitivas.

Un sistema de realidad aumentada puede definirse como aquél que enriquece el mundo real con diferentes tipos de elementos virtuales que son generados mediante técnicas hardware y software, permitiendo la coexistencia en un mismo espacio de objetos reales con objetos virtuales.

Azuma lo define como un sistema que reúne las siguientes características (Azuma et al, 2001):

- a) Combinación de imagen real y virtual.
- b) Interacción en tiempo real.
- c) Localización 3D.

En esta definición se pueden considerar sistemas de realidad aumentada tanto aquellos que utilicen dispositivos de visualización DMDs (Digital Micromirror Device), como otro tipo de tecnologías que empleen dispositivos basado en monitor y móviles (PDAs, Smartphone, etc.). El uso de estos tipos de unidades debe mantener las características esenciales de la realidad aumentada, la cual se centra en complementar la realidad en lugar de reemplazarla, consiguiendo de este modo un mayor realismo en todas nuestras acciones.

2.2.1. Realidad aumentada vs Realidad virtual

El término realidad aumentada (Caudell y Mizell, 1992) fue creado por Tom Caudell en 1992, no obstante sus orígenes podemos encontrarlos treinta años antes cuando el director de fotografía Morton Heiling diseñó un simulador de bicicletas en el que se combinaban imágenes, sonido, vibración y olfato (Azuma, 1997). Para definir la realidad aumentada y la realidad virtual suele utilizarse el esquema de continuidad de Milgram (Azuma et al., 2001). Así, en 1994 Milgram dio un paso importante hacia la consolidación del concepto de realidad aumentada, definiendo el continuo virtual como el espacio que abarcaba desde el entorno real al entorno virtual. Uno de los padres de esta

tecnología es Azuma, que la presenta como una variación de la realidad virtual (Leiva, Guevara y Rossi, 2012).

La realidad virtual es una tecnología que engloba un amplio conjunto de ideas, a partir de las cuales muchos investigadores han centrado sus trabajos. El término fue definido como un entorno tridimensional, interactivo y generado por ordenador en el cual se sumerge una persona (Aukstakalnis y Blatner. 1992; Izkara, Los Arcos y Maseda, 2005).

Los puntos claves de esta definición son tres: el primero hace referencia a que el entorno virtual es una escena tridimensional generada por ordenador, la cual requiere de un importante rendimiento gráfico para poder conseguir un nivel de realismo adecuado; en segundo lugar, el mundo virtual es interactivo, es decir, se necesitan respuestas en tiempo real por parte del sistema para tener la sensación de interactividad; el tercer y último punto consiste en que el usuario debe encontrarse inmerso en este mundo virtual, por tanto el sistema debe poder extraer del usuario distinta información como por ejemplo movimientos, provocando efectos que se representarán sobre la escena que se muestre en el dispositivo de visualización.

En cambio, la realidad aumentada consiste en ampliar la realidad percibida utilizando elementos de la realidad virtual procedentes de dispositivos tecnológicos concretos. Una diferencia importante entre realidad aumentada y realidad virtual es que la última suministra un entorno totalmente envolvente para el usuario, la visión y en algunos casos la audición y percepción deben estar controlados por el propio

sistema. Por el contrario, la realidad aumentada debe complementar el mundo real, siendo necesario que el usuario mantenga el sentido de presencia en dicho mundo. Las imágenes virtuales se mezclarán con la vista real para crear la imagen aumentada (Pasman et al, 2004), además los objetos virtuales generados por los sistemas deben estar alineados correctamente con el mundo real, a todo este proceso se le denomina registración. El ajuste de las imágenes debe ser lo más exacto posible en todo momento, incluso cuando el usuario se encuentre en movimiento. En la tabla 2.1 podemos estudiar las diferencias y similitudes más significativas de estos dos tipos de tecnologías (Izkara, Los Arcos y Maseda, 2005):

Características	Realidad Aumentada	Realidad Virtual
Inmersión virtual	Parcial	Total
Control de los sentidos	Parcial	Visión total, otros total/parcial
Presencia de objetos del mundo real	Sí	No
Tiempo real	Sí	Sí
Tridimensional	Sí	Sí

Tabla 2.1. Diferencias entre realidad virtual y realidad aumentada

Como hemos comentado anteriormente, Milgram describe una taxonomía en la que se detalla la relación existente entre realidad aumentada y realidad virtual (Milgram y Kishino, 1994).

El mundo real y un mundo totalmente virtual son los dos extremos de una línea continua (Figura 2.1), en cuyo punto medio encontramos lo que denomina Milgram como Realidad Mixta (RM). Los sistemas de

realidad aumentada (RA) están más cercanos al entorno real, siendo el mundo real complementado con datos generados por diferentes técnicas hardware y software. La virtualidad aumentada (VA) es un término creado por Milgram para identificar sistemas que son principalmente sintéticos, pero que agregan ciertas imágenes del mundo real como videos y texturas sobre objetos virtuales.

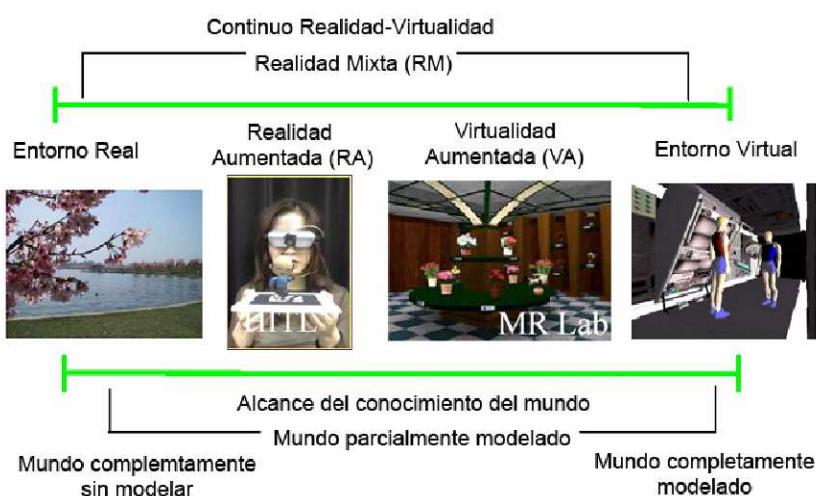


Figura 2.1. Relación entre realidad virtual y realidad aumentada

Según Milgram la realidad mixta se sustenta fundamentalmente en tres elementos (Milgram y Kishino, 1994; Holz et al, 2011):

- a) La fidelidad de la reproducción: el objetivo principal es la búsqueda del realismo en los dispositivos de realidad mixta, tomando como punto de referencia para ello la calidad de las imágenes.
- b) Alcance de la presencia metafórica: se encarga de medir el nivel de inmersión del usuario dentro de la escena visualizada. Cada

dispositivo que se utiliza tiene una visión distinta de la escena (Bimber y Raskar, 2005).

- c) Alcance del conocimiento del mundo: este tipo de técnicas no consisten simplemente en superponer un objeto virtual sobre el mundo real, es necesario mantener una perfecta relación entre los objetos virtuales y la imagen del mundo real. Por tanto, deben coexistir de una forma perfecta ambos tipos de imágenes. Esto requiere un detallado conocimiento de la relación entre el mundo real, la visión de la cámara y el usuario.

La realidad aumentada cambia, por tanto, la percepción del usuario del mundo real y la interacción con él, proporcionándole información que no puede detectar directamente por sus sentidos. Para poder conseguir esta percepción e interacción, se deberá utilizar dispositivos adicionales que permitan añadir información virtual a la información que percibe del mundo real.

De forma genérica, en un sistema de realidad aumentada es necesaria una viodecámara que capture la información del mundo real. Además debe poseer un sistema de geoposicionamiento que detecte la ubicación y orientación del usuario en cada momento, para que de esta forma pueda generarse la escena virtual que se integrará con la información del mundo real capturada. Esta escena aumentada debe presentarse a través de un dispositivo de visualización, estos sistemas se denominan *video-through* (Bimber y Raskar, 2005).

Los sistemas en los que la imagen real y virtual se mezclan en el cerebro del usuario, recibiendo solamente la imagen virtual a través del

dispositivo de visualización, se denominan *see-through* (Bimber y Raskar, 2005; Cawood y Fiala, 2008).

Existe otro tipo de realidad aumentada que consiste en proyectar la escena virtual sobre el entorno real, combinándose ambas en la propia escena real, en este caso se habla de sistemas basados en proyección.

2.2.2. Técnicas de inmersión de los usuarios

La realidad aumentada y la realidad virtual hacen uso de diferentes técnicas hardware y software para conseguir que el usuario se sienta presente en el entorno virtual generado. La descripción objetiva de ciertos aspectos del sistema como son la resolución de las pantallas, el campo de visión de las mismas, etc. se denomina inmersión (Slater y Wilbur, 1997). Las técnicas para lograr la inmersión del usuario se basan principalmente en la visualización sobre HMD (*Head-mounted display*) o CAVE (*Cave automatic virtual environment*).

Los HMD son los dispositivos que más se han utilizado tanto en realidad aumentada como realidad virtual (en esta técnica se basa el proyecto Google Glass). Existen dos tecnologías para mostrar gráficos sobre la vista del mundo real que tiene el usuario: la primera, denominada video see-through HMD, utiliza mezcla de video e imágenes en un HMD sin visión directa del mundo real total o parcial; mientras que la segunda, llamada optical see-through HMD, emplea combinaciones ópticas (esencialmente espejos semiplateados o pantallas LCD transparentes). La figura 2.2 muestra un esquema de ambas tecnologías.

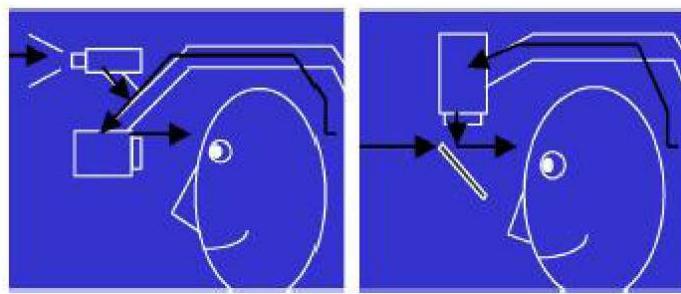


Figura 2.2. Video see-through y optical see-through³

La utilización de HMD presenta una serie de inconvenientes entre los que podemos destacar (Pérez, 2009):

- a) Falta de resolución.
- b) Campo de visión limitado.
- c) Ratio desequilibrado entre ópticas pesadas y dispositivos ergonómicos con una calidad de imagen baja.
- d) Aspectos relacionados con la percepción visual debido a la profundidad constante de la imagen.
- e) Los dispositivos optical see-through requieren una calibración difícil (depende del usuario y la sesión).
- f) Incrementan el índice de incomodidad especialmente durante movimientos rápidos de la cabeza.
- g) Los dispositivos optical see-through son incapaces de ofrecer efectos de oclusión.

Un sistema CAVE se puede definir como una habitación inmersiva, en la que uno de sus objetivos es reducir al máximo el hardware que lleva el usuario. Se trata de unas salas en las que se proyectarán imágenes desde el exterior a sus paredes, creando así un ambiente

³ (Pérez, 2009)

tridimensional donde se ubicarán los usuarios, que percibirán el mundo virtual creado a través de unas gafas especiales basadas en obturación, para separar los campos alternativos del video dirigido a cada ojo (Li y Xu, 2011).

Para provocar la inmersión en los entornos virtuales (CAVE o HMD) los estímulos visuales y auditivos deben estar siempre presentes, aunque a veces se pueden complementar con otros elementos como pueden ser estímulos táctiles.

Existen numerosos sistemas de realidad aumentada que requieren movilidad de los usuarios y necesitan acceder a la información en cualquier momento y en cualquier lugar, para ello se hace necesario el uso de dispositivos móviles y de software que sea capaz de ejecutarse en este tipo de dispositivos. Nos referimos a la realidad aumentada móvil (Visser, 2011), la cual ofrece importantes retos y dificultades añadidas a las presentadas en los entornos anteriormente descritos.

El rápido avance de las tecnologías y la miniaturización de los componentes electrónicos ofrecen un futuro prometedor a este tipo de sistemas, éste será el ámbito tecnológico en el que se desarrollará la tesis.

2.2.3. Geolocalización

Se trata de uno de los aspectos más importantes dentro de las tecnologías de realidad aumentada, ya que permite localizar la situación del usuario y posicionar en la pantalla adecuadamente las proyecciones virtuales (Thomas et al, 2000).

La geolocalización puede definirse como la capacidad de asignar coordenadas geográficas a diferentes elementos por medio de herramientas informáticas (Zoellner et al, 2009). Actualmente existe una generalización en el uso de la tecnología GPS en dispositivos como los teléfonos móviles y ordenadores personales, lo que ha permitido que esta capacidad esté al alcance de cualquier ciudadano (Leiva, Guevara y Rossi, 2012).

En el ámbito turístico ya existen algunas guías que utilizan la geolocalización para situar diferentes lugares de interés en un mapa. Una de las herramientas que emplea técnicas basadas en realidad aumentada más utilizada actualmente para la edición de contenidos es Layar⁴, que proporciona información adicional sobre nuestro entorno real.

Así, Layar combina la utilización del GPS, la brújula digital, la cámara y la conexión permanente en Internet. Así, la cámara captura imágenes de alrededor y las muestra en pantalla; el GPS determina la posición exacta y la brújula, la dirección en la que se está mirando. A partir de toda esta información, la aplicación toma los datos virtuales de Internet y los superpone en la imagen real de la pantalla del móvil. Layar tiene capas de información que ofrecen diversos contenidos.

La aplicación de la realidad aumentada al turismo no solo debe incluir lugares turísticos o de ocio, sino que también pueden resultar muy interesantes las referencias a las redes sociales. Algunos ejemplos podrían consistir en detectar si en un determinado sitio se encuentra

⁴ <http://www.layar.com>

alguno de nuestros amigos o si son sitios recomendados por algún conocido, así como recibir información sobre promociones que realizan algunos negocios, etc. Otra aplicación de la geolocalización es el espíritu competitivo que puede aparecer en los usuarios respecto a estimular el descubrimiento de lugares, ofrecer mejores consejos, etc.

2.3. Componentes y aplicaciones de un sistema de Realidad aumentada.

En esta sección vamos a analizar los componentes software y hardware que resultan necesarios para poder implantar un sistema de realidad aumentada. Analizaremos algunos sistemas existentes, identificando en cada caso los diferentes componentes hardware y software que necesitan los sistemas, así como la arquitectura en la que están basadas.

2.3.1. Componentes hardware

Cada sistema de realidad aumentada puede tener unas características particulares que hagan necesaria la existencia de algún componente especial, pero la mayoría requiere una serie de componentes comunes para poder llevar a cabo los diferentes procesos.

Además de los componentes comunes como unidades de procesamiento, almacenamiento, dispositivos de transmisión y recepción de datos, etc., serán necesarios otros tipos de dispositivos más específicos y relacionados con los sistemas de realidad aumentada. Entre ellos podemos destacar:

- Cámaras.

- Pantallas de visualización.
- Equipos holográficos 2D o 3D.

En la tabla 2.2 podemos ver un resumen de los diferentes componentes necesarios en un sistema de realidad aumentada (Azuma et al., 2001)

Técnica	Componente hardware
Captura <i>video-through</i>	Todo tipo de cámaras de video
Captura <i>see-through</i>	Cámaras integradas de video, HMD (<i>Head Mounted Display</i>)
Reconocimiento geo-posicionamiento	Antena GPS, sensores de movimiento
Reconocimiento híbrido	Antena GPS, sensores de movimiento, dispositivos de captura de imágenes
Almacenamiento	Disco duro convencional
Tratamiento imágenes	Procesadores con velocidad de procesamiento de imágenes
Comunicaciones locales	Tarjetas de red, conectores RJ45, antena WIFI, puntos de acceso
Comunicaciones móviles	Equipamiento GSM
Visualización <i>video-through</i>	Pantallas de video, monitores, proyectores
Visualización <i>see-through</i>	HMD, teléfonos móviles

Tabla 2.2. Componentes hardware de un sistema de realidad aumentada

2.3.2. Componentes software

El uso de un software viene determinado principalmente por las prestaciones y características del sistema de realidad aumentada que se desee implantar. Todo software que se utilice deberá tener un módulo de reconocimiento, ya sea de posicionamiento o por imágenes (con o sin marcadores). Si es necesario, deberán utilizarse librerías de

procesamiento de imágenes y software de comunicaciones. Por supuesto, el uso de las bases de datos, así como de las librerías de aumento de contenidos digitales serán necesarios en los sistemas de realidad aumentada (Seungjun y Anind, 2010). La mayoría de librerías de procesamiento de imágenes, como ARToolkit⁵, implementan tanto funciones de reconocimiento como de procesado y manipulación de contenidos visuales. En la tabla 2.3 (Azuma et al, 2001) podemos distinguir los diferentes requisitos software de un sistema de realidad aumentada:

Técnica	Componente software
Captura escena	Controladores de cámara de vídeo
Reconocimiento visual	Librerías de reconocimiento de imágenes
Reconocimiento geo-posicionamiento	Brújula digital, software GPS
Reconocimiento híbrido	Brújula digital, librerías de reconocimiento de imágenes, software GPS
Tratamiento de imágenes	Librería de tratamiento de imágenes
Almacenamiento	Base de datos
Comunicaciones locales	Controlador de red
Comunicaciones móviles	Controlador GSM
Visualización contenidos	Software de reproducción de contenidos multimedia, librerías de tratamiento de imágenes

Tabla 2.3. Requisitos software en un sistema de realidad aumentada

2.3.3. Aplicaciones existentes de Realidad aumentada

La realidad aumentada es una tecnología que tiene un gran futuro y un alto potencial por su aplicabilidad a un amplio conjunto de sectores.

⁵ <http://artoolkit.sourceforge.net>

Se han desarrollado aplicaciones experimentales en diversos entornos como industria del automóvil, medicina, marketing, arquitectura, construcción y educación.

En el campo de los videojuegos, la apuesta por la realidad aumentada es bastante fuerte para conseguir un gran número de adeptos. En el campo del transporte, la mayoría de las investigaciones apuestan por simulaciones de transportes aéreos y terrestres orientados hacia la mejora de los instrumentos de navegación. En el caso de emergencias sanitarias o similares, resulta muy interesante la línea para mejorar la evacuación de un determinado lugar; así como en cartografía, la creación de mapas interactivos. También pueden encontrarse proyectos de realidad aumentada relevantes en el ámbito educativo, médico y de la arquitectura.

Otra aplicación de la realidad aumentada en el ámbito de los transportes es su aplicación en los metros. Varias ciudades, como Madrid y Barcelona entre otras, tienen su propia aplicación de realidad aumentada que permite calcular los itinerarios que más le interesen a un turista, información sobre cada una de las líneas, frecuencia de paso, horarios, etc.

Otro ejemplo que podemos considerar son los dispositivos de visualización panorámica que se sitúan en miradores. Estos dispositivos no se han explotado como atractivo turístico porque el valor añadido que ofrece no es más que acercar los objetos que se encuentran alejados. Hoy en día, con la realidad aumentada se permite extender el concepto del visor panorámico al dispositivo de interpretación (Figura 2.3), el

cual con una apariencia similar puede incorporar elementos, para convertirlo en un dispositivo de realidad aumentada que sea capaz de complementar la información real. Un ejemplo comercial es InnoViewpoint⁶.



Figura 2.3. Visor panorámico

A rasgos generales, la realidad aumentada tiene que hacer frente a una serie de retos. La perspectiva temporal es necesaria para poder analizar adecuadamente su grado de implantación y éxito. Uno de los aspectos a mejorar en este tipo de tecnología son las técnicas relacionadas con los sensores de posición y orientación, porque aunque cada vez se va reduciendo el coste de estos dispositivos móviles, existe una oferta muy amplia y también una gran cantidad de diferentes dispositivos.

2.3.4. Arquitecturas de sistemas de Realidad aumentada

En la presente sección vamos a definir las diferentes arquitecturas en las que se sostienen los sistemas basados en esta tecnología, así como un

⁶ <http://www.innodevices.es/es/inicio/1.html>

análisis sobre el hardware y software utilizado por algunas propuestas representativas.

Los sistemas de realidad aumentada son sistemas informáticos en los que existe una importante interacción entre el sistema y el usuario, donde el tiempo de respuesta es un componente de especial importancia que determina en gran medida su utilidad.

Pueden encontrarse dos tipos de arquitecturas básicas en función de las prestaciones del sistema, de los componentes hardware y de la complejidad de procesamiento (López, Navarro y Relaño, 2010):

a) Sistemas autónomos: se trata de sistemas en los que la totalidad

de operaciones y tareas se llevan a cabo en un único terminal. Son aplicaciones que se ejecutan en entornos muy limitados en cuanto a espacio y toda la información debe estar almacenada en el terminal. La arquitectura interna de este tipo de aplicaciones suele estar compuesta por los siguientes módulos:

- Captación de escenario.
- Tratamiento de imágenes.
- Reconocimiento visual.
- Juego de patrones.
- Mezclado de realidad y aumento.
- Visualización.

b) Sistemas distribuidos: esta arquitectura delega en un equipo

servidor la tarea de realizar la mayor parte de los cálculos, rebajando de esta manera en los equipos de visualización la carga de proceso. En este tipo de sistemas, los terminales se encargan

de captar la escena y visualizar la realidad aumentada. El proceso intermedio lo realiza habitualmente el servidor, aunque en algunos casos varias tareas se pueden llevar a cabo en el mismo terminal para evitar sobrecargas en la red. No obstante, no suele ser lo común, por lo que la imagen captada por la cámara se envía completa o comprimida al servidor. Para el buen funcionamiento, es necesario disponer de dispositivos de envío y recepción de forma inalámbrica.

Entre las arquitecturas en sistemas distribuidos de realidad aumentada podemos destacar Ultra, March, Astor, Layar y Locus entre otros. El sistema Layar tiene una naturaleza que permite añadir contenidos, haciéndolo todo lo escalable que se deseé. Layar basa las tareas de reconocimiento por geoposicionamiento en el teléfono móvil, delegando la elección de la capa de superposición a los servidores dedicados. En las siguientes secciones vamos a describir las características técnicas de algunos de estos sistemas.

2.3.5. ULTRA⁷

Se trata de un proyecto cuyo objetivo era la creación de una plataforma de realidad aumentada ultra-ligera y compacta que permitiese a los desarrolladores la creación de sus propios sistemas. Este proyecto fue creado para que pudiera utilizarse fundamentalmente en dispositivos móviles, en concreto las PDAs (López, Navarro y Relaño, 2010).

⁷ <http://www.ist-ultra.org/>

El reconocimiento del escenario lo realiza utilizando marcadores, ya que este tipo de reconocimiento es más ligero, permitiendo aliviar la carga de trabajo y proceso al sistema. Entre las características necesarias del hardware del sistema de realidad aumentada se encuentran principalmente los aceleradores de vídeo y gráficos 3D. ULTRA utiliza como librería gráfica OpenGL.

Otro requisito importante de ULTRA es la necesidad de que el dispositivo tenga conexión W-LAN para poder comunicarse con el servidor, ya que permite la creación de contenidos off-line desde un servidor remoto. Para actualizar el sistema se hace necesario el trabajo de varios expertos para la creación de contenidos multimedia, imágenes 2D y 3D. En la tabla 2.4 destacamos las características técnicas de ULTRA.

Técnica	Componente software	Componente hardware
Reconocimiento marcadores	OpenGL ES	Cámara de teléfono móvil o PDA
Tratamiento imágenes	OpenGL ES	Procesador y memoria del teléfono móvil o PDA
Visualización	Delegado en software del dispositivo	Pantalla del teléfono móvil o PDA

Tabla 2.4. Características técnicas de ULTRA

2.3.6. MARCH

Se trata de un proyecto orientado a la visualización de distintos tipos de contenido digital sobre grabados de cuevas prehistóricas (Choudary et al., 2009; López, Navarro y Relaño, 2010). Al igual que

ULTRA este proyecto fue creado para que funcionara en dispositivos móviles.

Utiliza en el reconocimiento de escenas los marcadores. El uso de esta técnica se basa en poder emplear el sistema en cuevas reales y poder aligerar de carga de procesamiento y de cálculo al terminal.

El sistema consta de dos partes fundamentales:

- a) Elemento off-line que se encarga de registrar las imágenes que serán aumentadas en el terminal. Este proceso consiste fundamentalmente en añadir sobre la imagen del grabado información relevante por parte de los expertos.
- b) Aplicación móvil que se encargará, a partir del reconocimiento del marcador, de extraer la interpretación del experto.

Este modelo tiene un requisito fuertemente restrictivo, el uso de la librería gráfica NokiaCV que hace que solo los terminales Nokia puedan soportar la ejecución de esta aplicación. En la tabla 2.5 se resumen las distintas técnicas empleados en MARCH:

Técnica	Componente software	Componente hardware
Reconocimiento marcadores	Nokia CV	Cámara de teléfono móvil o PDA
Tratamiento imágenes	Nokia CVOpenGL ES	Procesador y memoria del teléfono móvil o PDA
Visualización	Delegado en software del dispositivo	Pantalla del teléfono móvil o PDA

Tabla 2.5. Características técnicas de MARCH

2.3.7. ASTOR

ASTOR es un prototipo de sistemas de realidad aumentada que está basado en que la información digital se muestra en unas pantallas transparentes, por lo que no es necesario que el usuario utilice ningún dispositivo (Olwal et al., 2005; López, Navarro y Relaño, 2010).

Para mostrar la información aumentada al usuario se emplea representación holográfica, utilizando para ello, proyectores que permiten reflejar la imagen sobre elementos holográficos ópticos (HOE) transparentes. ASTOR está pensado para trabajar en entornos pequeños, porque aunque podría usarse en entornos mayores, el coste que supondría sería muy alto. Todo el hardware se centra en el lado del proveedor y no en el del usuario, pues como hemos comentado anteriormente no necesitará ningún dispositivo. En la tabla 2.6 se resumen las características técnicas empleadas en este sistema:

Técnica	Componente software	Componente hardware
Reconocimiento imágenes	Java3DNokia CV	Cámara de vídeo
Tratamiento imágenes	Java 3D	NVIDIA GForce MX 440-SE
Comunicaciones	JRMI	Arquitectura cliente-servidor
Visualización	Delegado en software del dispositivo	Proyectores, pantallas HOE

Tabla 2.6. Características técnicas de ASTOR

2.3.8. POLAR

Se trata de un sistema de realidad aumentada de bajo coste que permite superponer información a una escena real utilizando para ello

unos mecanismos poco comunes (Owal y Höllerer, 2005; López, Navarro y Relaño, 2010). Hace uso de un espejo que refleja por un lado los objetos reales que se desean ampliar y por el otro lado, la información digital que se desea mostrar jugando con los diferentes ángulos de visión del espejo. Además, se utiliza una pantalla convencional desde la que se genera la imagen virtual reflejada en el espejo, una cámara y un dispositivo de detección de distancia que percibirá el movimiento de la vista del usuario, para que de esta forma se cambie el enfoque. En la figura 2.4 se muestra el funcionamiento del sistema POLAR.

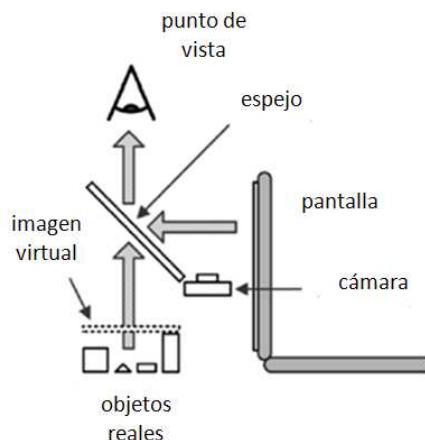


Figura 2.4. Funcionamiento del sistema POLAR

En la tabla 2.7 observamos el resumen de técnicas empleadas:

Técnica	Componente software	Componente hardware
Reconocimiento imágenes	OpenGL ES	WebCam, sensor de movimiento
Tratamiento imágenes	OpenGL ESJava 3D	Procesador y memoria del PC
Visualización	Delegado en software del dispositivo	Pantalla del PC, espejo medio reflectante

Tabla 2.7. Características técnicas de POLAR

2.3.9. Layar

Es una aplicación que permite la visualización de información digital sobre un entorno real utilizando dispositivos móviles. En todo momento el terminal del usuario va mostrando la imagen real que recoge su cámara.

Mediante el GPS del terminal se envía a un servidor la localización del usuario junto a la orientación extraída con la brújula del mismo. Esta información es de vital importancia, para que el servidor responda con la creación de una capa con la imagen o texto del lugar que se está enfocando con el móvil. La figura 2.5 muestra un esquema de la arquitectura Layar.

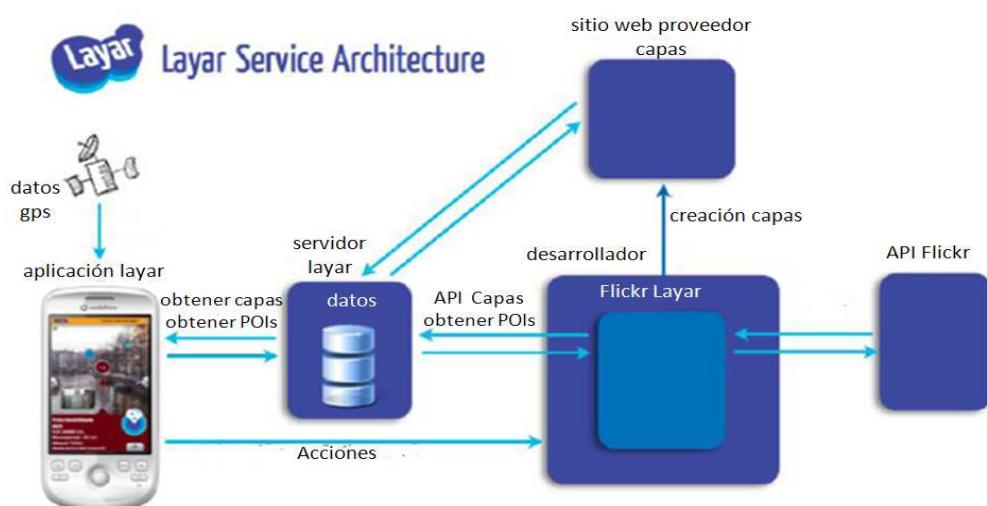


Figura 2.5. Funcionamiento del sistema Layar

Los requisitos que tiene el uso de esta aplicación son:

- Disponer de un teléfono móvil que utilice Android o iOS.
- Móvil con cámara de video.
- Móvil con antena GPS.

Estos requisitos serían para el usuario, mientras que para poder suministrar este servicio es necesario disponer de un servidor que reciba las peticiones de los distintos terminales y que les envíe la información que deberá mostrarse en el móvil. Es indispensable la utilización de una base de datos con todas las localizaciones y las informaciones a mostrar en las capas, lo que requiere un personal que se encargue de actualizar y mantener los datos. En la tabla 2.8 resumimos las técnicas empleadas en Layar.

Técnica	Componente software	Componente hardware
Reconocimiento geo-posicionamiento	Brújula digital del dispositivo, software GPS	Cámara del teléfono móvil o PDA, antena GPS
Almacenamiento de imágenes	Gestor de base de datos	Servidor de base de datos
Visualización	Delegado en software del dispositivo	Pantalla del teléfono móvil o PDA

Tabla 2.8. Características técnicas de LAYAR

Layar es un auténtico navegador para realidad aumentada, lo que significa que nos proporciona información adicional sobre nuestro entorno real como enlaces a wikipedia, datos sobre hoteles, bares, restaurantes, etc. Su nombre viene de layer, es decir, capa en inglés, ya que lo que hace el navegador es poner sobre la imagen real distintas capas de información extraídas de la red.

2.3.10. LOCUS⁸

Es un sistema de realidad aumentada desarrollado para terminales móviles. El usuario puede ver información ampliada sobre determinados lugares con su dispositivo gracias a las técnicas de geo-posicionamiento del sistema (López, Navarro y Relaño, 2010). Para hacerlo posible utiliza modelos 3D para aumentar la escena real. Además, crea herramientas con las que puede caracterizarse el contexto espacial en el que los usuarios van a requerir realidad aumentada.

Los requisitos del sistema son los siguientes:

- a) Almacenamiento y gestión de la información espacial y posicional de los lugares.
- b) Herramientas de modelado con las que realizar imágenes en 3D.
- c) Herramientas de navegación y estimación de rutas.
- d) Interfaz para trabajar con realidad virtual y realidad aumentada.

Técnica	Componente software	Componente hardware
Reconocimiento geo-posicionamiento	Brújula digital del dispositivo, software GPS	Cámara del teléfono móvil o PDA, antena GPS
Visualización	Delegado en software del dispositivo	Pantalla del teléfono móvil o PDA

Tabla 2.9. Características técnicas de LOCUS

⁸ www.locus.org.uk

2.4. Componentes de un navegador de Realidad aumentada

Los navegadores de realidad aumentada son un software que combina categorías, capas de información y otros elementos de diferente naturaleza, que son dispuestos sobre una pantalla donde se visualiza la realidad capturada por una cámara en tiempo real, permitiendo al usuario distinguir e interactuar con dichos elementos.

Este tipo de herramienta muestra la información cercana sobre una vista captada por la cámara del dispositivo. La aplicación calcula la posición del usuario en base a sus coordenadas, la brújula digital y el acelerómetro, accediendo a distintas bases de datos desde las que descarga información geográfica y otros datos de diversos puntos de interés.

En la figura 2.6 se muestran los diferentes elementos que contienen todos los navegadores de realidad aumentada.

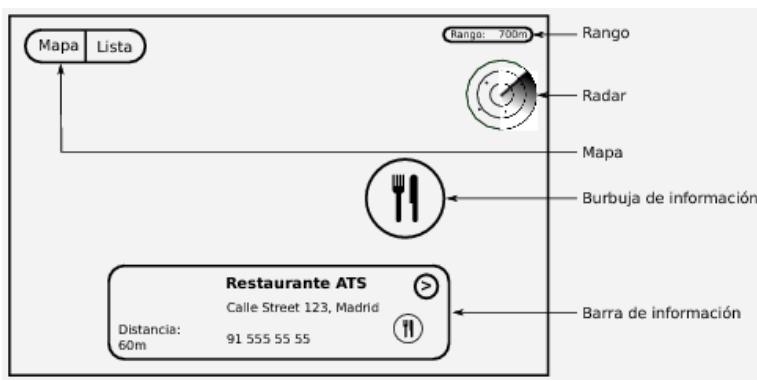


Figura 2.6. Esquema de elementos comunes a los navegadores de RA

El radar tiene como función mostrar la posición de los puntos cercanos al lugar donde se encuentra el usuario. Los diferentes puntos de interés se mostrarán sobre la pantalla mediante las burbujas de información, que normalmente contienen un ícono lo suficientemente descriptivo como para que sea reconocido rápidamente (en la figura 2.6 se muestra el ejemplo de un restaurante y aparece reflejado mediante un símbolo fácilmente reconocible). Cuando el usuario pulsa sobre alguna de las burbujas de información, debe desplegarse una barra de información que contenga datos adicionales sobre el punto (en el ejemplo aparece el nombre del restaurante, su dirección, su teléfono y distancia). El componente rango tiene como función indicar el radio alrededor del usuario en el que se buscan puntos. Normalmente es un parámetro configurable, aunque normalmente no es adecuado que sea demasiado grande. El elemento mapa permite pasar al modo de vista de mapa, que sitúa los diferentes puntos detectados sobre un mapa, permitiendo en ocasiones obtener información para llegar fácilmente hasta ellos. El elemento lista tiene la misma función, pero en forma de lista de elementos.

Algunos ejemplos de navegadores de realidad aumentada son:

- a) ARViewer SDK es un *framework* de realidad aumentada para dispositivos Android creado por el grupo de investigación GSyC/Libresoft de la Universidad Rey Juan Carlos I. ARViewer ha sido diseñado partiendo de la modularización del navegador de realidad aumentada LibreGeoSocial⁹. Incluye diferentes módulos

⁹ <http://www.libregeosocial.org/>

que permiten etiquetar puntos de interés y visualizarlos. Entre sus principales características destacan: (i) es capaz de visualizar información multimedia como imágenes, vídeo, sonidos y notas de texto; (ii) proporciona una interfaz de realidad aumentada que tiene en cuenta latitud, longitud y altitud; (iii) permite un modo de búsqueda; (iv) es capaz de mostrar cajas con información detallada de los elementos y (v) posee capacidad para etiquetar puntos a diferentes alturas. ARViewer puede funcionar en los siguientes modos:

- Independiente. Como aplicación que muestra puntos de interés obtenidos de diferentes fuentes de datos.
- Invocado desde otra aplicación.
- Versión modificada. Como única aplicación basada en ARViewer.

El componente que actúa como navegador utiliza las coordenadas de los puntos de interés y del usuario. Tiene la posibilidad de utilizar el factor de altitud, lo que permite trabajar con diferentes etiquetas en la misma geoposición situadas a distinta altura. Además, soporta la localización en interiores mediante el uso de códigos QR. Las etiquetas asociadas a los puntos pueden mostrarse tanto en modo de lista como en modo vista de navegador.

- b) Kharma¹⁰: es un proyecto de Georgia Institute of Technology, también basado en la filosofía del software libre. Una de las principales características de Kharma radica en la posibilidad de

¹⁰ <https://research.cc.gatech.edu/polaris>

utilizar herramientas de desarrollo web convencional (HTML o Javascript) para crear e integrar fácilmente contenidos. Otro punto a destacar es la precisión en la localización de contenidos digitales, para ello utiliza unos servicios denominados Infrastructure Services¹¹ y GEOSpot Services¹². Se trata de sistemas de georreferencia que ubica al usuario con mucha mayor precisión de la alcanzada por los sistemas GPS. Esta plataforma también permite a los desarrolladores crear y alojar contenido, utilizando para ello una versión extendida de un lenguaje de codificación llamado Keyhole Markup Language (KML), que actualmente es utilizado en Google Earth. El navegador Kharma permite a los usuarios abrir múltiples canales con contenido de realidad aumentada, cada uno de ellos añadiendo contenido visual a la escena y permitiendo que puedan interactuar entre sí, utilizando para ello protocolos web estándar.

- c) Blippar¹³: se trata de una aplicación gratuita para Smartphone dirigida a anunciantes. En esta App el contenido depende de cada marca. Con esta aplicación podríamos por ejemplo ver cómo nos quedaría un reloj en nuestra muñeca, etc.
- d) Otros navegadores de realidad aumentada muy conocidos son Layar, Wikitude, Mixare y Junaio, que están orientados sobre todo al guiado de los usuarios. En la siguiente sección se harán referencias a ellos.

¹¹ <https://research.cc.gatech.edu/polaris/content/infrastructure-service>

¹² <https://research.cc.gatech.edu/polaris/content/geospot-service>

¹³ <http://blippar.com>

2.5. Aplicaciones de la Realidad Aumentada al turismo

La Realidad Aumentada permite una gran interacción y la posibilidad de mostrar información de una forma intuitiva, rápida y atractiva. Este es uno de los principales motivos por el cual este tipo de tecnología se está aplicando en el turismo, patrimonio, cultura y publicidad. Son sectores que requieren que sus productos resulten atractivos para el público, así como que ofrezcan una fácil interactividad.

Por otra parte, cada vez más los turistas demandan un mayor valor añadido a sus viajes y visitas (Leiva, Guevara y Rossi, 2012), más allá de las “clásicas” guías y mapas en papel impreso, audio-guías, videos informativos o consultas a través de un ordenador, que no ofrecen apenas posibilidad de interacción ni de personalizar la información disponible (Figura 2.7).



Figura 2.7. Audio guía, Cortometraje y TouchTech Tourist Guide

Es muy habitual encontrar en museos y centros de información turística el uso de pantallas táctiles interactivas (Figura 2.7), como la touchtech tourist guide, para ofrecer información más personalizada e interactiva. Todos estos soportes, habituales actualmente, cumplen las necesidades básicas de información de los turistas; sin embargo, se sigue

demandando información más interactiva, personalizada y fácil de manejar en cualquier momento y lugar.

Las empresas privadas y administraciones públicas están comenzando a realizar importantes inversiones en la utilización de las nuevas tecnologías, para realizar una promoción de los recursos patrimoniales y turísticos. Esta apuesta no sólo responderá a las necesidades actuales, sino que es un importante activo para convertirse en un futuro próximo en un elemento atractivo y diferenciador para los visitantes.

En estos últimos años se están desarrollando una importante cantidad de herramientas para dispositivos móviles basados en realidad aumentada, las cuales pueden tener aplicaciones al turismo. Se pueden distinguir dos importantes grupos (Leiva, Guevara y Rossi, 2012):

- Aplicaciones y proyectos turísticos orientados al patrimonio. Se trata de las primeras aplicaciones que surgieron y que tienen como principal objetivo la reconstrucción, animación o visualización virtual (normalmente mediante el uso de modelos 3D más o menos complejos) de monumentos y edificios emblemáticos que se encuentran actualmente en ruinas. Estas aplicaciones tienen la característica de que suelen requerir dispositivos más o menos especializados y/o potentes por el coste computacional que supone la representación de modelos 3D (Caro, 2012).
- Aplicaciones y proyectos turísticos orientados al guiado. En este grupo se encuentran, en general, el resto de aplicaciones. Suelen

estar enfocadas a guiar u orientar al turista en un recorrido o ciudad, indicándole dónde se encuentra un determinado recurso turístico. Este tipo de aplicaciones son más genéricas que las anteriores y en general suponen un menor consumo de recursos, dado que no tiene la necesidad de visualizar modelos en 3D.

Estos dos tipos de aplicaciones solucionan las demandas del sector turístico para conseguir una mejor experiencia por parte del turista en la visita del destino. Resulta muy interesante la sinergia de ambos tipos de aplicaciones, permitiendo de esta forma conseguir por parte del turista una mayor comprensión de los monumentos, el ambiente y la evolución histórica de los sitios visitados, consiguiendo que la inmersión del turista sea mayor y más gratificante.

Otro gran problema que encuentran muchos turistas al visitar un destino es averiguar dónde se localizan los sitios. Es muy habitual perderse y desorientarse, aun a pesar de la utilización habitual por parte del turista de mapas; pues no todas las personas son capaces de orientarse satisfactoriamente y a veces, las indicaciones pueden cambiar mucho de un mapa a la realidad por obras o desactualización.

Por tanto, el uso de aplicaciones de Realidad Aumentada permitirá solucionar en gran medida estos problemas, además de aportar soluciones personalizadas para cada turista. Este tipo de aplicaciones, además de ser de gran utilidad, son muy usables en el turismo debido a que la mayoría de las mismas son (o tienden a ser) portátiles, lo cual favorece la ubicuidad y el acceso a la información en cualquier punto,

permitiendo una mayor, más fácil y natural interactividad entre el usuario y el sistema.

2.5.1. Aplicaciones y proyectos turísticos orientados al patrimonio

En este grupo se engloban las aplicaciones y proyectos cuyo principal objetivo es la reconstrucción, animación o visualización virtual de monumentos y edificios emblemáticos. El objetivo final de estos proyectos es permitir difundir el patrimonio monumental del destino, así como el aprovechamiento de las tecnologías de la información como elemento atractivo para potenciar el turismo cultural.

Dentro de este conjunto de aplicaciones podemos incluir ejemplos como el proyecto Lifeplus que utiliza la Realidad Aumentada para realizar una representación de Pompeya (Vlahakis, Demiris y Ioannidis, 2004), en la que el usuario puede interactuar con personajes ficticios y observar los monumentos de la ciudad, hoy en día en ruinas, reconstruidos de forma virtual (Figura 2.8). Este proyecto presenta al visitante una percepción del entorno en tiempo real de granrealismo, ofreciendo una gran inmersión y una experiencia única. Otro proyecto similar, también de la misma empresa (Miralab), es el proyecto Virtual Hagia Sophia, cuyo objetivo es la reconstrucción virtual de la iglesia Hagia Sophia de Estambul (Foni, Papagiannakis y Magnenat-Thalmann, 2002).



Figura 2.8. Proyecto Lifeplus

Un tercer proyecto de aplicación orientada a la reconstrucción virtual arquitectónica es Archeoguide (Vlahakis et al., 2002), el cual consiste en crear un sistema para visualizar modelos 3D que se superponen a monumentos reales (Figura 2.9), de los cuales apenas quedan los cimientos. Los modelos 3D utilizados son sencillos y utilizan una resolución gráfica baja para de esta forma facilitar que el procesamiento sea rápido y el sistema pueda funcionar en tiempo real.



Figura 2.9. Representación virtual del Templo de Hera

Además de los proyectos cuyo objetivo es la reconstrucción virtual de monumentos, la realidad aumentada puede también ser utilizada para mostrar otro tipo de información. Un ejemplo de ello es el museo Miraikan de Tokio, que emplea Realidad Aumentada para ofrecer información sobre las piezas del museo de forma virtual superponiendo la información junto a la figura, de forma que no es necesario añadir

carteles o explicaciones físicas junto a las piezas. Para conseguir este efecto, se utilizan unas gafas especiales que son proporcionadas a los visitantes del museo.

Todos los ejemplos comentados son de comienzos del 2000 y en su mayoría requieren de dispositivos específicos (HMD, receptor de GPS, etc.); pero presentan el inconveniente de que son sistemas voluminosos y aunque sean portátiles, resultan incómodos y de difícil difusión, debido a la necesaria inversión de los museos o instituciones para tener una gran cantidad de sistemas en funcionamiento.

Actualmente, la mayoría de los proyectos hacen uso de dispositivos móviles de amplia difusión, como son smartphones, tablets, pdas o portátiles ultraligeros, que están prácticamente al alcance de cualquiera e incluyen normalmente accesorios como cámara, conexión a Internet y GPS. Se intenta realizar un aprovechamiento de la tecnología de la información en el turismo cultural y hacer más accesible el patrimonio monumental y artístico. En el proyecto iTacitus (Zoellner, et al, 2009) no se renderizan¹⁴ modelos 3D, pero sí imágenes y fotografías mediante el lenguaje para gráficos vectoriales X3D, utilizando como plataforma un PC Ultra Móvil (UMPC) o un dispositivo móvil con conectividad a Internet. En concreto lo que hace es un filtrado de la realidad, cambiando el modo de vista a un modo en escala de grises que permite ver los edificios como eran originalmente a través del UMPC (Figura 2.10). Otra aplicación interesante es Google Goggles que es un servicio

¹⁴ Término informático para referirse al proceso de generar una imagen o video partiendo de un modelo 3D.

de Google que permite reconocer objetos mediante fotos tomadas con un dispositivo móvil, devolviendo información relacionada.



Figura 2.10. Representación del Palacio de Diana (Italia)

2.5.2. Aplicaciones turísticas orientadas a guiado

En esta sección vamos a describir las características de las aplicaciones que tienen como objetivo el guiado de los turistas en un destino turístico o ciudad, indicando dónde se hallan los monumentos de interés, restaurantes, hoteles y otros productos turísticos. Esta idea se aplicó mediante Realidad Aumentada hace diez años (Wagner y Schmalstieg, 2003).

Estas aplicaciones habitualmente necesitan utilizar una serie de dispositivos que poseen una amplia difusión, como pueden ser tablets, móviles o smartphones (Figura 2.11), en lugar de utilizar dispositivos específicos como gafas de Realidad Aumentada, HMD, etc.



Figura 2.11. Dispositivos para aplicaciones orientadas al guiado

Obviamente, la ventaja de usar este tipo de plataformas es que los sistemas desarrollados tendrán un público potencial sensiblemente mayor que otros basados en un hardware específico o poco accesible. Además, la utilización de dispositivos habitualmente conocidos por el usuario facilita la interacción, la ubicuidad y la personalización de la información. Aparte de todas estas ventajas, habrá que tener en cuenta el hecho de que la mayoría de estos dispositivos incluyen accesorios como cámara, conexión wifi y GPS, que hacen que dispongamos de un dispositivo de mayor potencialidad y una gran versatilidad. No obstante, una de las debilidades de utilizar este tipo de dispositivos es que su capacidad de cómputo es reducida.

Existen en el mercado varios ejemplos de guías virtuales, entre ellas se encuentra Wikitude¹⁵, que se trata de una de las aplicaciones pioneras en su ámbito y que ha dado una gran difusión a la Realidad Aumentada (Figura 2.12). Concretamente, Wikitude es un navegador de Realidad Aumentada desarrollado por la empresa austriaca Mobilizy GmbH para iOS y Android y publicado en 2008 bajo licencia freeware.

El funcionamiento de Wikitude se basa en mostrar sobre la propia vista de la cámara del móvil información relevante como hoteles, monumentos, etc. Para saber qué puntos de interés mostrar, la aplicación hace uso del GPS o redes WiFi para determinar la posición del móvil, del compás para obtener la orientación (dónde está enfocando la cámara) y los acelerómetros para estimar la altura a la que se está enfocando respecto al nivel del mar. Basándose en esta información, la

¹⁵ <http://www.wikitude.com>

aplicación permite mostrar puntos de interés, que contendrán una descripción del mismo, un enlace Web, un teléfono, etc. Dichos puntos son agrupados en categorías llamadas Mundos Wikitude (Wikitude Worlds en inglés) y cada usuario decide qué categorías habilitar y por tanto visualizar. El contenido que muestra en cada punto es generalmente proporcionado por los usuarios y puede ser generado o modificado usando KML (*Keyhole Markup Language*) o ARML (*Augmented Reality Markup Language*) (Visser, 2011).



Figura 2.12. Uso de Wikitude en un smartphone

Posteriormente han ido surgiendo diferentes navegadores de realidad aumentada. Uno que ha adquirido gran fama y que ya hemos comentado es Layar (Figura 2.13), que está desarrollado por la empresa SPRX Mobile y fue lanzado al mercado en 2009 para iOS y para Android. Al igual que Wikitude, Layar hace uso del GPS, compás y acelerómetros, para determinar la posición y orientación del móvil y con esta información superponer a la imagen de la cámara del móvil

distintos puntos de interés. En el caso de Layar, los puntos están agrupados en lo que se denominan *capas* (Layers en inglés). Cuando se selecciona un punto de interés, el navegador es capaz de mostrar imágenes, reproducir audio, vídeos, llamar a un número de teléfono, lanzar aplicaciones, etc. Al igual que Wikitude, las capas pueden ser generadas por cada usuario y hospedadas en el servidor de Layar para su uso.



Figura 2.13. Uso de Layar en un smartphone

Se han diseñado diversas aplicaciones turísticas basadas en Layar (Costa del Sol Occidental, Guía Turística de Segovia, Camino de Santiago,...), todas ellas han sido creadas mediante capas. Son gratuitas y debido a que funcionan con Layar, únicamente se encuentran disponibles para los dispositivos iOS y Android.

Permiten crear filtros para poder visitar el destino sin tener que seguir los típicos itinerarios de todas las guías y así potenciar la enorme cantidad de recursos de los que se disponen.

En el caso de la guía turística de Segovia estos filtros son: iglesias, museos, palacios, judería, obra civil y eventos. De esta forma se pueden localizar los diferentes edificios importantes de la ciudad con información sobre cómo llegar, posibilidad de audio guías, acceso a la web para preguntar precios u horarios, etc.

Otro ejemplo de aplicación basada en Layar es *Costa del Sol Occidental*, que ha sido desarrollada por el consorcio Qualifica. Ofrece información turística de los municipios de la Costa del Sol Occidental (Torremolinos, Benalmádena, Mijas, Fuengirola, Marbella, Estepona, Casares y Manilva), permitiendo consultar la información sobre los recursos desde la web y también desde la aplicación descargable para los dispositivos Android e iOS. Es una herramienta creada para potenciar el turismo de la zona, incidiendo positivamente en la competitividad y la innovación. Se trata de un sistema de información con más de 5.000 puntos de interés turísticos y permite seleccionar categorías como parques de ocio, playas, turismo activo, alojamientos y congresos en un rango de búsqueda de hasta 70km.

Junaio¹⁶ es otro navegador de Realidad Aumentada que ha sido desarrollado por la empresa Metaio GmbH para iOS y Android. Lanzado al mercado en 2009, la principal diferencia de este navegador con respecto a Wikitude y Layar es que, además de usar GPS para detectar la localización del móvil, Junaio también hace uso de la triangulación de antenas para detección en interiores

¹⁶ <http://www.junaio.com>

Al igual que las otras dos aplicaciones, Junaio permite crear canales propios y habilitar sólo aquellos en los que se está interesado. Además, permite dejar comentarios y opiniones de sitios concretos que otros usuarios pueden leer. Sin embargo, la principal diferencia respecto a las anteriormente descritas es la capacidad de reconocimiento de imágenes, códigos QR y de barras.

La tecnología denominada Junaio Glue asocia información a cualquier objeto identificado, lo que permite que se pueda añadir información al reconocer un cuadro, una entrada de un restaurante, etc. La información que Junaio puede asociar a los puntos de interés es, al igual que en Layar, imágenes, audio, vídeo, modelos 3D estáticos y además modelos 3D animados (Figura 2.14).

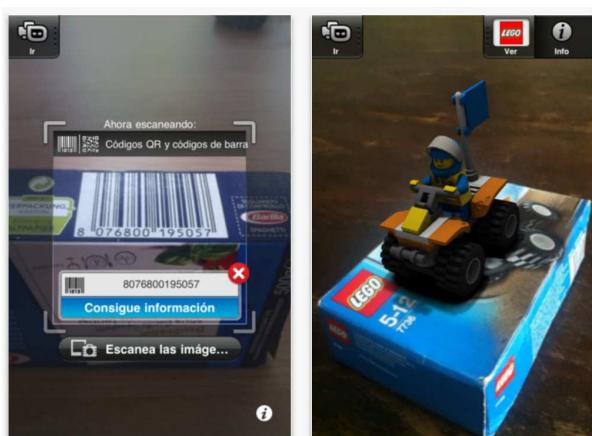


Figura 2.14. Uso de Junaio en funcionamiento

Los ejemplos comentados anteriormente son quizás los más conocidos y más utilizados, pero existen otros menos utilizados como MARA¹⁷ (navegador de RA desarrollado por Nokia), Enkin¹⁸ (navegador de RA

¹⁷ <http://research.nokia.com/page/219>

¹⁸ <http://www.enkin.net>

para Android), Toozla¹⁹ (navegador de Realidad Aumentada basado en audio), ARO²⁰ (navegador de RA similar a Wikitude), Yelp²¹ (para búsqueda de restaurantes, salones de belleza, bares, ...) y GUIMO²², una guía virtual multimedia de La Garriga (ciudad de la comarca del Vallés Oriental en la provincia de Barcelona) que proporciona información interactiva sobre los puntos de interés turístico, de ocio y comercial de la ciudad. Se basa al igual que Wikitude y Layar en la utilización de GPS para su localización y funcionamiento. Otro ejemplo sería Mixare²³, un motor de Realidad Aumentada open source bajo licencia GPL y navegador desarrollado por la empresa Peer Internet Solutions (Figura 2.15). Este motor, al igual que Wikitude y Layar, se basa en GPS para la localización y permite añadir contenidos o usar datos de la wikipedia, google, twitter, etc.



Figura 2.15. Uso de Mixare en funcionamiento

Como puede observarse, son muchas las aplicaciones donde la Realidad Aumentada puede ser aplicada al sector turístico para:

- Conseguir una mayor y mejor difusión del patrimonio y la cultura. Su utilización en museos, edificios históricos y puntos de

¹⁹ <http://www.toozla.com>

²⁰ <http://www.orizom.com/HtmlContenido.aspx?IdHtmlContent=61>

²¹ <http://www.yelp.es>

²² <http://www.visitlagarriga.cat/es/guimo>

²³ <http://www.mixare.org/>

interés turístico puede resultar de gran utilidad para mostrar información sobre estos recursos de forma atractiva, interactiva, divertida y fácil de entender. Esto será un valor añadido a tener en cuenta para los turistas.

- Ofrecer información personalizada, interactiva y adecuada a la situación y momento. Mediante su uso es posible encontrar comentarios de visitantes sobre si merece la pena visitar un museo, qué restaurante de los más cercanos es el mejor y más económico o simplemente información sobre un edificio desconocido y que no aparece en las guías y mapas clásicos.

Siendo un gran avance y unas herramientas con un gran presente y futuro, destacamos que uno de los grandes inconvenientes de todos estos navegadores de realidad aumentada es la no utilización de atributos contextuales, ni de sistemas de recomendación individuales/grupales que permitan al usuario ahorrar tiempo localizando y buscando aquello que le interese.

2.5.3. Estado de la técnica en repositorios de contenidos

Los contenidos que las aplicaciones de Realidad Aumentada muestran son muy variados y podemos distinguir los siguientes:

- Texto. Se trata de uno de los recursos de contenido más utilizados para aumentar la realidad en multitud de sistemas y aplicaciones. Las principales ventajas de utilizar esta técnica son que apenas requiere procesamiento y ocupa poco ancho de banda para su recepción, ya que la mayoría de las aplicaciones de realidad aumentada acceden a repositorios externos para

recuperar y acceder al contenido aumentado que muestran. Además, existe una gran cantidad de repositorios accesibles, como wikipedia, twitter o repositorios personalizados (como los existentes para Wikitude o Layar) que permiten añadir comentarios, evaluaciones, etc. La mayor desventaja del texto es la dependencia del idioma (aunque hoy en día no suele ser un gran problema, ya que lo habitual es que la aplicación pueda trabajar en diferentes idiomas) y que no es muy atractivo visualmente. Así, por ejemplo, el texto es usado en aplicaciones que no tienen mucha potencia de cálculo o no tienen gran ancho de banda, siendo un contenido ideal para aplicaciones basadas en móviles y dispositivos portátiles como son Wikitude, Layar, Junaio o Mixare, ya que el texto es un contenido básico para casi cualquier aplicación (Figura 2.16).



Figura 2.16. Uso de contenido textual en realidad aumentada

- Imágenes. Se trata de otro de los contenidos más utilizados habitualmente para aumentar la realidad. Las ventajas de las imágenes son que en general, aunque depende de sus dimensiones, requieren poco ancho de banda y poco procesamiento para su visualización, aunque mayor que el texto. Sin embargo, es innegable que las imágenes son bastante más llamativas y

atractivas. Además, pueden ser mostradas como iconos, fotografías, reconstrucción virtual de monumentos o modelos 3D.

- Audio. Es otro de los posibles contenidos que pueden existir en una aplicación de Realidad Aumentada. Consumo más ancho de banda y de recursos que las imágenes para su reproducción, aunque no suele ser usado como elemento principal de transmisión de información (excepto en el caso de audio guías o guías para ciegos), sino que se utiliza normalmente para ambientar, pues normalmente en entornos ruidosos el audio puede no ser muy efectivo.
- Vídeo. Este tipo de contenido también puede ser utilizado en aplicaciones de Realidad Aumentada. La ventaja de este medio es que es muy atractivo, pero tiene un alto costo computacional y de ancho de banda.
- Objetos 3D. Estos a su vez pueden ser estáticos o animados. Lo normal es que sean estáticos, aunque actualmente la tendencia es animarlos, sobre todo para publicidad (Figura 2.17).



Figura 2.17. Objeto 3D de realidad aumentada en publicidad

A continuación se expone en la tabla 2.10 un resumen de las características más relevantes de los diferentes contenidos que pueden mostrarse en aplicaciones de realidad aumentada.

Contenido	Ancho de banda necesario	Coste computacional	Atractivo
Texto	Bajo	Bajo	Bajo
Imágenes	Bajo/Medio	Bajo/Medio	Medio
Audio	Medio	Medio	Medio/Bajo
Vídeo	Medio/Alto	Medio/Alto	Medio/Alto
Objetos 3D	Alto	Alto	Alto

Tabla 2.10. Características de contenidos en aplicaciones de RA

2.5.4. Plataformas de repositorios de contenidos.

Existen diferentes repositorios de contenidos donde poder extraer distintos tipos de información que puedan mostrarse en una aplicación de Realidad Aumentada. Se agrupan en:

- Repositorios de texto. Existen diferentes tipos de repositorios de información textual, siendo los más habituales:
 - Repositorios de texto geoposicionados. Incluyen texto y datos asociados a una ubicación geográfica. Muestran la información dependiendo de la posición del usuario. Entre los repositorios de texto geoposicionados más interesantes y típicos encontramos google maps y bing maps. Disponen de sus propias APIs (Application Programming Interface) que permiten acceder al contenido disponible en dichos repositorios.

- Repositorios generados por usuarios. El uso de estos tipos de repositorios de texto generados por usuarios es algo bastante común en la mayoría de aplicaciones de Realidad Aumentada de guía virtual (Wikitude, Layar, Junaio...). Concretamente, el repositorio más usado es la Wikipedia, una enciclopedia libre en la que cualquier usuario puede editar el contenido. Este repositorio es muy utilizado, ya que el contenido de la Wikipedia se puede usar libremente y dispone de una API que permite el acceso de programas clientes al contenido de ésta. Además, hay otros repositorios de contenidos generados por usuarios, como los mapas públicos de google maps o bing maps, capas de Layar, etc.
- Repositorios de redes sociales. Las redes sociales como Facebook o Twitter pueden considerarse como otro tipo de repositorio textual. Existen navegadores de Realidad Aumentada cuyo único cometido es mostrar la información de redes sociales de las personas cercanas a la posición del usuario, como TwittAround, por ejemplo.
- Repositorios de búsquedas generales. En esta categoría se incluyen los buscadores como google, bing, yahoo, etc., que mediante sus APIs permiten realizar búsquedas relacionadas con los puntos de interés cercanos (obtenidos de los repositorios de textos geoposicionados).
- Repositorios de imágenes. En el caso de las imágenes la clasificación se divide en repositorios de pago y gratuitos, siendo

estos últimos generalmente del estado o bien generado por los usuarios.

- **Repositorios de pago.** Existen muchos bancos de imágenes disponibles de pago, entre los que destacan Shutterstock images²⁴, Fotolia²⁵ o Bigstock²⁶. Algunos bancos de imágenes de pago permiten un uso ilimitado y otros que el pago se realiza cada vez que se use una imagen.
- **Repositorios gratuitos.** Podemos incluir aquellos bancos de imágenes que tienen un uso educacional y no comercial. Entre ellos destacamos el banco de imágenes del Ministerio de Educación²⁷, así como repositorios generados por usuarios, como Panoramio²⁸, Flickr²⁹ o Picassa³⁰. Estos bancos de imágenes tienen sus propias APIs de acceso. En general, la licencia de uso es para aplicaciones no comerciales, aunque por ejemplo la API de Panoramio es gratuita para uso comercial. Otra ventaja de utilizar repositorios como Panoramio, Flickr o Picassa es que las imágenes están geoposicionadas, lo que facilita su uso para navegadores de Realidad Aumentada.
- **Repositorios de audio.** Al igual que en el caso de los repositorios de imágenes se clasifican en gratuitos y de pago.

²⁴ <http://www.shutterstock.com>

²⁵ <http://es.fotolia.com>

²⁶ <http://www.bigstockphoto.es>

²⁷ <http://recursostic.educacion.es/bancoimagenes/web>

²⁸ <http://www.panoramio.com>

²⁹ <http://www.flickr.com>

³⁰ <http://picasa.google.com>

- Repositorios de pago. No existen tantos como en el caso de imágenes, ya que hay más repositorios de sonidos gratuitos que de pago. Un ejemplo de repositorio de sonidos de pago sería SoundSnap³¹ o iStockphoto³².
 - Repositorios gratuitos. En este tipo de repositorios se incluyen los pertenecientes a organismos públicos y repositorios generados por usuarios. Respecto a los primeros, está el repositorio de sonidos del Ministerio de Educación que ofrece sus contenidos para uso no comercial; de entre los repositorios generados por usuarios, está Flashkit³³, MediaCollege³⁴ o FindSounds³⁵. Otra opción es el uso de Podcasts, que consiste en la distribución de archivos multimedia (audio o vídeo) normalmente mediante un sistema de redistribución (RSS). La mayoría de Podcasts son de libre uso.
- Repositorios de vídeo. Al igual que en los dos casos anteriores, se clasifican también en repositorios de pago y gratuitos.
 - Repositorios de pago. Entre ellos podemos destacar iStockphoto³⁶, Shutterstock³⁷ o Pond5³⁸.
 - Repositorios gratuitos. Un ejemplo es el repositorio de vídeos pertenecientes al Ministerio de Educación y

³¹ <http://www.soundsnap.com>

³² <http://www.istockphoto.com/audio>

³³ <http://www.flashkit.com/>

³⁴ <http://www.mediacollege.com/downloads/sound-effects/>

³⁵ <http://www.findsounds.com/>

³⁶ <http://www.istockphoto.com/video>

³⁷ <http://footage.shutterstock.com/>

³⁸ <http://www.pond5.com/>

también los repositorios generados por usuarios, como youtube³⁹, metacafe⁴⁰, o yahoo vídeos⁴¹.

- Repositorios de objetos 3D. También se clasifican en gratuitos y de pago.
 - Repositorios de pago. Este tipo de repositorios requieren el pago de cada modelo para poder ser descargado y según el caso tienen uso ilimitado o no. Los más conocidos son: kit3Dmodels⁴², Mr.Cad⁴³ y geo-metricks⁴⁴.
 - Repositorios gratuitos. Respecto a repositorios de modelos 3D gratuitos, hay páginas para descargar gratuitamente, como arq.com⁴⁵ o 3DCadModel⁴⁶. También existen páginas de intercambios de modelos 3D como 3DModelSharing⁴⁷ y también existen repositorios públicos como Sketch Up⁴⁸.

2.6. Análisis DAFO de la Realidad aumentada aplicada al Turismo

A continuación se expone un análisis DAFO (Debilidades, Amenazas, Fortalezas y Oportunidades) de la realidad aumentada en el sector turístico (Troitiño, Rodriguez y Hidalgo, 2011; Leiva, Guevara y Rossi, 2012):

³⁹ <http://www.youtube.com/>

⁴⁰ <http://www.metacafe.com/>

⁴¹ <http://es.video.yahoo.com/>

⁴² <http://www.kit3dmodels.com/>

⁴³ <http://www.mr-cad.com/>

⁴⁴ <http://www.geo-metricks.com/>

⁴⁵ http://documentos.arq.com.mx/Planos_y_Modelos/Modelos_3D/index.html

⁴⁶ <http://www.3dcadmodel.com/>

⁴⁷ <http://www.3dmodelsharing.com/>

⁴⁸ <http://sketchup.google.com/3dwarehouse/>

2.6.1. Debilidades

Implantar esta tecnología en el sector turístico conlleva una serie de debilidades que enumeramos a continuación:

- a) El sector turístico es especialmente sensible a la coyuntura económica actual y no es un buen momento para la inversión en nuevas herramientas y métodos.
- b) Necesidad de cambiar los hábitos de la oferta y la demanda para introducir tecnologías de esta naturaleza.
- c) La realidad aumentada está orientada a un segmento de mercado muy específico, por tanto es necesario que tienda hacia la accesibilidad universal.
- d) Se debe disponer de forma permanente de medios para realizar actualizaciones de contenidos.
- e) Ofrece un número excesivo de puntos de interés, lo que dificulta su usabilidad.

2.6.2. Amenazas

Respecto a las amenazas que tiene la implantación de esta tecnología podemos indicar:

- a) Una falta de equilibrio entre las políticas de precios de las compañías telefónicas y la demanda de la realidad aumentada. Esto se hace mucho más crítico cuando el usuario quiere utilizar sistemas de realidad aumentada en países diferentes al suyo. Es necesario que las compañías telefónicas implanten tarifas más económicas para satisfacer la demanda de este tipo de tecnología.
- b) Se necesita tener todas las bases de datos constantemente actualizadas, lo cual requiere un elevado coste.

- c) Una fuerte competitividad dentro del sector.
- d) Elevados costes de inversión inicial y de actualización.

2.6.3. Fortalezas

No obstante, a pesar de las debilidades y amenazas, existen un gran número de fortalezas que hacen que se consoliden los sistemas de tecnología basado en realidad aumentada en el sector turístico.

- a) Aumenta la experiencia turística del visitante mejorando y ampliando la propia realidad del producto turístico, convirtiéndola en única.
- b) Permite al turista combinar la información virtual con datos reales de manera interactiva y atractiva para el visitante.
- c) Puede permitir al usuario decidir los contenidos que le interesan de acuerdo a diferentes tipos de criterios (tipo de contenido, tiempo, motivaciones, etc.).
- d) Es una línea totalmente innovadora en el ámbito turístico, que permite integrarlo en un marco de gestión de proyecto cultural de una zona, ciudad, etc.
- e) Se trata de una apuesta totalmente directa del turismo por las nuevas tecnologías y por tanto de un producto de calidad.
- f) Amplía la utilización de dispositivos móviles con conexión de datos a Internet, lo que facilita la transmisión de información una vez generada la aplicación.
- g) Independencia del turista para gestionar su viaje.

2.6.4. Oportunidades

Entre las oportunidades que tienen los sistemas de realidad aumentada aplicada al turismo destacamos:

- a) Permite al usuario decidir los contenidos que le interesan de acuerdo a diferentes criterios.
- b) Es una línea innovadora en el ámbito turístico y resulta interesante integrarlo en un marco de gestión completa de un destino.
- c) Es un elemento diferenciador dentro del competitivo mercado turístico que busca la excepcionalidad y la calidad.
- d) Permite la creación de proyectos multidisciplinares, que integran mediante una estrategia de actuación, diferentes ámbitos como el patrimonio cultural, el turismo, urbanismo, etc.
- e) Aprovechar la experiencia en otros campos, para trasladarla y aplicarla al sector turístico.
- f) Se trata de un soporte aplicable a diversos componentes del sistema turístico, potenciando la trazabilidad.
- g) Permite la combinación con otro tipo de herramientas como pueden ser los Sistemas de Información Geográficos (SIG).

Capítulo 3

Sistemas de Recomendación

3.1. Introducción a los sistemas de recomendación

Las herramientas de realidad aumentada aplicadas al turismo que están surgiendo en los últimos años ofrecen una gran variedad de posibilidades al turista. La cantidad de información que puede ofrecernos es tan grande y tan completa en muchos de los casos, que podemos considerar la aplicación de este tipo de técnicas como una importante fuente de diversidad en la oferta turística (Leiva, Guevara y Rossi, 2012).

Uno de los problemas que podemos encontrarnos es que las personas o usuarios finales tendrán que manejar de forma efectiva una importante cantidad de datos, teniendo muchas veces una gran sobrecarga de información.

Este inconveniente provoca que los turistas empleen mucho tiempo para realizar una búsqueda efectiva sobre la información total. Esto puede dar lugar a que los beneficios potenciales de la realidad aumentada se vean minimizados, a causa de la cantidad de información que el turista tendrá que ir desechar y el gran número de resultados que obtiene. Por tanto, deben establecerse mecanismos para que los resultados que le aparezcan al usuario sean adecuados respecto a sus expectativas iniciales y que no tenga que interactuar en demasía con el sistema.

En los últimos años, el uso de sistemas de recomendación (SR) se ha popularizado en aplicaciones de muy diverso tipo (iTunes Store, Amazon, etc.). En el ámbito del turismo, resulta bastante interesante su utilización en aplicaciones web que faciliten la planificación de viajes, ofreciendo una selección personalizada de productos. Este tipo de sistemas se hace necesario en los destinos turísticos, con objeto de que el turista pueda encontrar fácil y rápidamente productos adecuados a sus gustos, ya que lo habitual es que la web del destino ofrezca una amplia lista de puntos de interés. La no utilización de sistemas de recomendación, termina provocando en muchos casos que los usuarios se sientan desbordados al tener que elegir entre una gran diversidad de alternativas para obtener los resultados esperados (Lymberopoulos et al., 2011, García, Torre y Linaza, 2014). Por tanto, los SR deben implementar mecanismos de filtrado para proporcionar un conjunto de puntos de interés que se ajusten de la forma más precisa posible a las necesidades reales del turista.

Los sistemas de recomendación son un mecanismo para poder recuperar resultados que tienen una amplia utilización en motores de búsqueda como por ejemplo Google. Como su nombre indica, este tipo de sistemas se encargan de recomendar o sugerir a los usuarios que lo utilizan ítems relacionados a sus preferencias (Burke, 2000).

Un sistema de realidad aumentada aplicado al turismo tiene que ser sensible a las preferencias del turista, debiendo establecer mecanismos que permitan recomendarle posibles ofertas relacionadas con sus preferencias e intereses. Debe existir una estructura con los gustos concretos para favorecer el incremento de la lealtad al sistema.

Actualmente, los turistas tienen en Internet una fuente de información inagotable en la que poder consultar información sobre los puntos turísticos de interés del destino que van a visitar o que están visitando. La mayoría de los destinos estimaron que la red es un entorno fantástico para poder acceder a los potenciales turistas, por lo que el número de sitios web con información de los destinos, sus recursos, ofertas, etc., han crecido de tal forma que actualmente los turistas se encuentran muchas veces desbordados con tanta información.

Adaptar la información a los gustos de los turistas o de grupos de personas que viajan juntos parece una necesidad que los destinos turísticos tendrían que afrontar y de esta forma adaptar los sistemas web construidos. La realidad es diferente y los turistas se encuentran con mucha información que en el mejor de los casos pueden filtrar por categorías establecidas por el destino.

Para solucionar este problema de sobrecarga de información, se han planteado diferentes métodos que pretenden mejorar la satisfacción de los turistas. La personalización es una herramienta que permite adaptar un servicio a los gustos y preferencias particulares de quien solicita la búsqueda de información.

Un caso muy utilizado es la personalización web (Lawrence et al, 2000; Pérez, 2008), que ofrece contenidos adaptados y una presentación de las páginas basándose en las preferencias del usuario obtenidas a partir de visitas anteriores.

La personalización consiste en ofrecer información de los productos adecuados de forma precisa a los usuarios que los demanda. Las herramientas que proporciona este tipo de facilidades se denominan sistemas de recomendación (Burke, 2000; Mooney y Roy, 2000; Schwab, Kobsa y Koychev, 2001; Burke, 2002; Niu et al, 2002; Mak, Koprinska y Poon, 2003; Pérez, 2008, Leiva, Guevara y Rossi, 2012).

Existen diferentes tipos de situaciones donde el uso de sistemas de recomendación es muy útil y necesario, como por ejemplo:

- a) Recomendación en tiendas on-line: partiendo de un producto se recomiendan otros productos que han interesado a otros usuarios que compraron dicho producto. Un ejemplo pionero fue la web Amazon.com
- b) Filtrado de noticias: se construye un perfil que almacena las noticias que un usuario consulta.
- c) Recomendaciones musicales o de cine: cada vez que un usuario escucha por ejemplo una canción, se envía esta información a la

base de datos del sistema, el cual las utiliza para generar nuevas recomendaciones.

- d) Búsqueda de personas afines en comunidades: en webs como menéame.net se tienen en cuenta las noticias que cada usuario ha votado para generar una lista de vecinos con similares intereses.

Los sistemas de recomendación fueron definidos inicialmente como aquellos sistemas que, partiendo de recomendaciones proporcionadas por los propios usuarios, las añadía y las mostraba a usuarios apropiados (Resnick y Varian, 1997). Actualmente el término sistema de recomendación es mucho más amplio, siendo una herramienta muy atractiva en situaciones donde el conjunto de información es muy grande y resulta ineficiente por parte de los usuarios la búsqueda de información (Leiva, Guevara y Rossi, 2012).

Estos tipos de sistemas deben permitir generar tanto recomendaciones individuales como grupales, teniendo como objetivo el guiar de forma personalizada a aquellos elementos que resultarán más interesantes dentro de un amplio abanico de opciones. Actualmente muchos sitios web han comenzado a utilizar sistemas de recomendación (Leiva, Guevara y Rossi, 2012).

En el presente capítulo vamos a realizar un amplio estudio sobre los diferentes modelos de recomendación existentes, centrándonos principalmente en aquellos que vamos a utilizar en nuestro trabajo.

3.2. Mecanismos de Recomendación

Las técnicas de recomendación tienen un amplio abanico de aplicaciones. Así, podemos encontrar herramientas de comercio electrónico para la venta de productos, utilidades para recomendar la lectura de libros, asistentes para ayudar a decidir qué actividades turísticas visitar, etc.

Vamos a centrarnos en el ámbito turístico. Para recomendar a un turista la visita de una determinada actividad o punto de interés de un destino, son aplicables varios tipos de técnicas. Pueden utilizarse criterios relacionados con los puntos más visitados en el destino turístico, tipo de visitas que realiza habitualmente el visitante del destino, información demográfica de los turistas, etc.

También los resultados de las recomendaciones al turista pueden mostrarse de diferente forma, aunque lo habitual es una lista de actividades para que el turista seleccione y pueda obtener información detallada, ver opiniones de otros usuarios, etc.

El uso de sistemas de recomendación por parte de los turistas de un destino potenciará la fidelización, si las recomendaciones generadas son buenas. De esta forma, mediante la interacción con los sistemas por parte del turista se conseguirá aprender sus demandas y gustos. Si el proceso de personalización tiene éxito, existen muchas posibilidades de que el turista vuelva a utilizar el sistema y esto hará que se obtenga más información personal del turista sobre sus preferencias y así tendremos unos clientes más fieles al sistema.

En esta sección vamos a presentar una clasificación de sistemas de recomendación (Resnick y Varian, 1997; Terveen y Hill, 2002; Pérez, 2008; Leiva, Guevara y Rossi, 2012), para comprender mejor la clasificación, previamente debemos conocer qué elementos deben contener:

- a) Datos de campo. Este tipo de información es el conjunto de datos que el sistema debe poseer antes de que el proceso de recomendación comience.
- b) Datos de entrada. Es el conjunto de datos que el usuario comunica al sistema para generar una recomendación. Incluye la información relacionada con las preferencias, gustos y necesidades del usuario.
- c) Algoritmos para realizar la recomendación. Deben combinar los datos de campo con los datos de entrada para generar las recomendaciones solicitadas.

La información de entrada es esencial para que los algoritmos de recomendación generen los resultados adecuados. El objetivo es sugerir nuevos elementos a un usuario basándose en sus elecciones anteriores y en las realizadas por otros usuarios con historial similar de calificaciones. Dependiendo de la forma de obtener dicha información podemos distinguir entre:

- a) Información explícita (Cho y Kim, 2004; Herlocker et al, 2004; Adomavicius y Tuzhilin, 2005). El usuario informa al sistema de sus preferencias respecto de los ítems. La forma de realizarlo normalmente es mediante la asignación de una puntuación a cada

elemento, que será habitualmente un valor numérico discreto entre un máximo y un mínimo. Se trata de información que proporciona el propio usuario a partir de peticiones que el sistema le solicita. Normalmente al darse de alta en el sistema, será interrogado mediante formularios y cuestionarios para conocer sus gustos y preferencias. El usuario es responsable de la veracidad de la información aportada, por lo que en gran medida los resultados que posteriormente generará el sistema de recomendación dependerá de esta información.

- b) Información implícita (Adomavicius y Tuzhilin, 2005; Pérez, 2008; Leiva, Guevara y Rossi, 2012). Se recoge información a partir de las acciones del usuario (tiempo que pasa leyendo una información determinada, enlaces que sigue, número de veces que escucha una canción...), mediante una aproximación clásica de minería de datos. De esta forma, se registran las acciones de los usuarios y se concluyen las calificaciones sobre estas acciones. El sistema, a partir de la interacción realizada con el usuario, recogerá la información de forma automática. Normalmente esta información puede obtenerse a partir de datos reflejados en el historial de navegación, visitas realizadas anteriormente, etc.

Antes de comenzar a describir la clasificación de los sistemas de recomendación, realizaremos algunas definiciones que nos ayudarán a comprender mejor los diferentes tipos de sistemas existentes:

Definición 3.1. Sea I el conjunto de ítems sobre los cuales se realizan recomendaciones.

Definición 3.2. Sea U el conjunto de usuarios de los que se tiene conocimiento de sus gustos y sus preferencias.

Definición 3.3. Sea u el usuario sobre el que se va a realizar una recomendación.

Definición 3.4. Sea i un ítem del que se desea realizar una predicción sobre el grado de interés que tendría su visita para el usuario u .

Aunque podemos encontrar diferentes clasificaciones de sistemas de recomendación, podemos distinguir en la mayoría de la literatura los siguientes tipos, en función de las técnicas que utilizan:

- a) Sistemas de recomendación colaborativos (Billsus y Pazzani, 2000; Sarwar et al, 2001; Schwab, Kobsa y Koychev, 2001; Pérez, 2008, Maleszka, Mianowska y Nguyen, 2013). La mayoría de los sistemas de recomendación utilizan esta técnica y posiblemente sean los más extendidos en el mercado. Este tipo de técnica añade valoraciones o recomendaciones a los objetos, identificando los gustos comunes de los usuarios basándose en sus valoraciones para posteriormente generar una nueva recomendación teniendo en cuenta las comparaciones entre ellos. Por tanto, los resultados que proporcionan se basan en las valoraciones realizadas por usuarios con gustos similares al que realiza la petición de recomendación. En un sistema de recomendación colaborativo un perfil de usuario es un vector de objetos y sus valoraciones. La principal ventaja es que son independientes de la representación de los productos que se recomiendan. En la tabla 3.1 se resumen sus características principales.

Datos de campo	Valoración por parte de todos los usuarios pertenecientes a U de los diferentes ítems contenidos en I .
Datos de entrada	Valoración por parte de u de los ítems contenidos en I .
Algoritmo de recomendación	Identifica usuarios en U que tenga preferencias similares a u , prediciendo la valoración de un ítem i a partir de las valoraciones de ese grupo de usuarios.

Tabla 3.1. Técnica de recomendación colaborativa

b) Sistemas de recomendación basados en contenido (Schwab, Kobsa y Koychev, 2001; Mak, Koprinska y Poon, 2003; Bezerra y Carvalho, 2004; Pérez, 2008; Leiva, Guevara y Rossi, 2012). Se basan en categorizar los ítems a recomendar, proporcionando resultados que tengan características similares a otros que han sido valorados anteriormente por el usuario. Cada objeto a recomendar está asociado y definido por sus características. El sistema debe aprender un perfil de intereses de los usuarios, basándose en las características presentes de los ítems que el usuario selecciona. En la tabla 3.2 se resumen sus características principales:

Datos de campo	Datos sobre los atributos y características de los ítems pertenecientes a I .
Datos de entrada	Valoración por parte de u de los ítems contenidos en I .
Algoritmo de recomendación	Identifica el comportamiento de las valoraciones realizadas por u y genera una clasificación de aquellos ítems pertenecientes a I que cumplen con esos atributos y características.

Tabla 3.2. Técnica de recomendación basada en contenido

c) Sistemas de recomendación demográficos (Vozalis y Margaritis, 2007; Pérez, 2008; Leiva, Guevara y Rossi, 2012). Clasifica a los usuarios según diferentes parámetros personales y realiza las recomendaciones teniendo en cuenta el grupo demográfico al que pertenece el usuario. La representación demográfica de la información puede variar enormemente del sistema y del tipo de productos que vayan a recomendarse. Por tanto, se basa en la utilización de estereotipos de usuarios. Uno de los principales beneficios de los sistemas de recomendación demográficos es que no requieren información histórica, requisito que es necesario en los sistemas colaborativos y basados en contenido. El principal inconveniente de esta técnica es requerir información demográfica sobre el usuario (muchas veces de carácter personal), lo que hace

que muchos puedan sentir vulnerada su privacidad. En la tabla 3.3 se resumen sus características principales:

Datos de campo	Información demográfica sobre U y la valoración de los diferentes ítems contenidos en I .
Datos de entrada	Información demográfica de u .
Algoritmo de recomendación	Identifica usuarios de U que son demográficamente similares a u y a partir de las valoraciones de estos usuarios realiza una predicción de la valoración que tendría para u el ítem i .

Tabla 3.3. Técnica de recomendación demográfica

- d) Sistemas de recomendación basados en conocimiento (Burke, 2002; Pérez, 2008). Este tipo de sistemas dispone de información sobre cómo un ítem satisface a un usuario y establece una relación entre necesidad y recomendación. Las recomendaciones se basan en la inferencia sobre las necesidades y las preferencias del usuario. Aunque todos los métodos de recomendación realmente realizan inferencias, este tipo de técnica se distingue del resto en que usa un conocimiento funcional, ya que partiendo de la información de cómo un objeto en particular puede satisfacer las necesidades del usuario, razona la relación entre éstas y una posible recomendación. El perfil del usuario es una

estructura de conocimiento que apoya la inferencia. En la tabla 3.4 se resumen sus características principales:

Datos de campo	Información basada en las características de los diferentes ítems de I e información sobre cómo estos ítems satisfacen las necesidades de u .
Datos de entrada	Información sobre las necesidades, demandas e intereses de u .
Algoritmo de recomendación	Infiere la relación entre i y las demandas de u .

Tabla 3.4. Técnica de recomendación basada en conocimiento

- e) Sistemas de recomendación contextuales (Adomavicius et al, 2010; Leiva, Guevara y Rossi, 2012). Los sistemas de recomendación descritos en los apartados anteriores utilizan exclusivamente información del perfil del usuario o de los objetos a recomendar, no dándole importancia a la información contextual. La relevancia de este tipo de información ha sido reconocida por investigadores y profesionales de muchas disciplinas (Lamsfus et al., 2009; Miele et al., 2013; Mettouris y Papadopoulos, 2014), incluyendo aquellos que se dedican a la recuperación de información, minería de datos, etc. El uso de información contextual tales como hora, lugar o la compañía de otras personas (por ejemplo, para ver películas o salir a cenar) es de vital importancia en la prestación de las recomendaciones e

indudablemente para las turísticas. El uso de la información contextual puede realizarse en diferentes paradigmas algorítmicos, siendo un nuevo elemento a tener en cuenta en los futuros sistemas de recomendación.

- f) Sistemas de recomendación híbridos (Burke, 2002; Leiva, Guevara y Rossi, 2012). Todas las técnicas descritas brevemente en los apartados anteriores presentan limitaciones y problemas. Para solucionar estos inconvenientes, se plantea la posibilidad de utilizar técnicas híbridas de recomendación. Podríamos considerar sistemas de recomendación híbridos a aquellos que utilizan dos o más técnicas de recomendación con el objetivo de mejorar el rendimiento de los resultados obtenidos.
- g) Sistemas de recomendación grupales (Jameson y Smyth, 2007, Leiva, Guevara y Rossi, 2012). Tradicionalmente los sistemas de recomendación han sido diseñados para usuarios individuales. Resulta muy interesante la posibilidad de desarrollar técnicas que permitan proponer recomendaciones a grupos de usuarios que viajan juntos. Este tipo de técnicas plantea varios problemas, fundamentalmente relacionados con la forma de adquirir las preferencias del grupo. Para la adquisición de información, la mayoría de los sistemas grupales utilizan técnicas similares a las aplicadas en los sistemas individuales, complementándola con información sobre las relaciones entre los diferentes miembros del grupo.

Una limitación muy habitual en las diferentes técnicas de recomendación, es el problema de incremento (ramp-up problema) (Konstan et al., 1998), que se presenta en las siguientes situaciones:

- a) No disponer de datos históricos del usuario puede hacer que el sistema no funcione adecuadamente. Así, cuando un usuario se registra en el sistema y las recomendaciones únicamente tienen en cuenta valoraciones anteriores o usuarios ‘supuestamente’ similares al que solicita la recomendación, los resultados no serán óptimos. Además, en algunos sistemas cuando existen pocas valoraciones por parte de un usuario, es posible que se le asigne un grupo inadecuado.
- b) Cuando un nuevo ítem se da de alta en el sistema, al no tener valoraciones de él, difícilmente puede ser recomendado.

Otro tipo de problema que puede surgir es cuando se cambia radicalmente las preferencias el usuario. En estos casos, muchos sistemas de recomendación solucionan este problema mediante la inclusión del envejecimiento de valoraciones, de tal manera que las antiguas posean menor influencia. Esta solución plantea el riesgo de perder información valiosa para las preferencias del usuario.

Para concluir este punto podemos afirmar que el turismo es una actividad donde las preferencias e intereses del turista son fundamentales. Para que un sistema de recomendación orientado al turismo cumpla adecuadamente su función debe contemplar los siguientes factores:

- Mantenimiento de una estructura que almacene los gustos y preferencias del turista.
- Considerar si se viaja en grupo las preferencias de todos los miembros.
- Registro de las valoraciones que los turistas otorgan a los puntos de interés visitados.
- Debe tenerse en cuenta no sólo las preferencias de los turistas, sino también la información contextual (Lamsfus et al, 2009).

En las siguientes secciones revisaremos con más profundidad los sistemas de recomendación enumerados.

3.2.1. Sistemas de recomendación colaborativos

El término minería de datos es un campo de las ciencias de la computación que surge cuando se realizan procesos de análisis de grandes volúmenes de datos, para descubrir patrones y comportamientos. El principal objetivo del proceso de minería de datos consiste en extraer información de un conjunto de datos, para poder transformarla en una estructura comprensible para un uso posterior (Eirinaki y Vazirgiannis, 203; Cho y Kim, 2004; Pérez, 2008).

Los primeros sistemas de recomendación colaborativos basaban sus esquemas de funcionamiento en técnicas de minería de datos. Normalmente seguían dos fases: una primera, dedicada a conocer y aprender el modelo; una segunda, en la cual aplican el modelo obtenido en la fase anterior a situaciones reales, generando las recomendaciones. Sin embargo, resulta mucho más eficiente que el modelo pueda ir adaptándose durante el funcionamiento, ya que las bases de datos

evolucionan dinámicamente a partir de las interacciones de los usuarios con el sistema.

En muchas ocasiones los ítems que un usuario demanda en un sistema de búsqueda de información son muy parecidos a los utilizados por otros usuarios con intereses parecidos. Por tanto, pueden aprovecharse las búsquedas realizadas por unos usuarios para facilitar la tarea a otros con gustos similares a él. Un sistema de recomendación colaborativo recomienda ítems no por su contenido, sino porque hay usuarios similares al solicitante que evaluaron positivamente esos ítems. De esta manera, no se analiza la similitud entre ítems sino la similitud entre usuarios.

El problema de este tipo de sistemas aparece cuando un nuevo ítem se incorpora al sistema. Hasta que un usuario no lo evalúe no se dispone de ninguna información para poder recomendarlo, pues no se lleva a cabo ningún análisis de su contenido. Otro problema importante es cuando el sistema posee pocos usuarios, ya que existe una baja probabilidad que evalúen un mismo ítem y si lo hacen, que la evaluación sea similar. En este caso, el sistema sería bastante ineficiente (Bueno, 2002).

Los sistemas de recomendación colaborativos pueden ser clasificados en dos grupos (Breese, Heckerman y Kadie, 1998; Deshpande y Karypis, 2004; Jin, Chai y Si, 2004; Zeng et ak,m 2004; Adomavicius y Tuzhilin, 2005; Schafer et al., 2007):

- a) Basados en modelos (Breese, Heckerman y Kadie, 1998; Adomaviciu y Tuzhilin, 2005; Pérez, 2008): estos algoritmos,

antes de realizar una recomendación, deben desarrollar primero un modelo basado en las calificaciones de los usuarios. El proceso de producción de los modelos que usan los algoritmos en esta categoría se desarrollan con distintos enfoques de aprendizaje automático como Redes Bayesianas, encasillamiento o enfoques basados en reglas, etc.

- b) Basados en memoria (Resnick y Varian, 1997; Bueno, 2002): utilizan la base de datos completa de usuarios-ítems para generar las predicciones. El sistema de recomendación usa una técnica estadística para hallar un conjunto de usuarios que tengan una historia de concordancia con el usuario objetivo (vecinos). Una vez formada una vecindad, se utiliza un algoritmo que combine las preferencias de los vecinos, para producir una predicción y un ranking de los “N principales” ítems a recomendar al usuario objetivo.

Los sistemas de recomendación colaborativos dependen de las coincidencias en las calificaciones de los diferentes usuarios, pero cuando el número de valoraciones a un mismo ítem es escaso tiene dificultades en encontrar este tipo de coincidencias. A este inconveniente se le suele denominar densidad del espacio de valoraciones. Este problema de densidad se reduce cuando se aplican enfoques basados en modelos. También aparece este hándicap cuando el número de ítems a recomendar es muy grande y por tanto las coincidencias son escasas.

Si la cantidad de ítems es cambiante, las valoraciones existentes apenas tienen influencia para los nuevos usuarios, que no tendrán

calificaciones para comparar. Si el número de ítems es muy grande y se tienen pocas valoraciones, entonces las probabilidades de que sean comunes también serán muy escasas.

Por tanto, podemos concluir que un sistema de recomendación que aplique técnicas colaborativas tendrá un buen funcionamiento, cuando la densidad de usuarios sea alta en comparación con el número de ítems, que debe ser pequeño y estático (Pérez, 2008).

El hecho de tener usuarios asignados a grupos con gustos similares, mejorará la eficiencia de los sistemas de recomendación colaborativos. Cuando un usuario no encaja en ningún grupo, esto hará disminuir también la eficiencia del sistema. Se trata de una característica que se presenta muy habitualmente en los sistemas de recomendación demográficos, aunque este tipo de sistema no tiene el problema de “nuevo usuario”, ya que no requieren sus valoraciones, pero tiene el inconveniente de tener que recopilar información demográfica.

3.2.1.1. Basado en modelos.

A continuación haremos una breve exposición sobre los diferentes algoritmos que implementan un filtrado colaborativo basado en modelos para la obtención de recomendaciones, destacando entre ellos:

- a) Redes bayesianas de creencia (Breese, Heckerman y Kadie, 1998):
se trata de una red gráfica que representan relaciones probabilísticas entre variables. Razonan bajo incertidumbre y combinan las ventajas de una representación visual intuitiva con una base matemática basada en la probabilidad bayesiana. Las redes bayesianas son muy prácticas en situaciones donde el

conocimiento sobre el usuario cambia muy lentamente, siendo por tanto ineficiente en aquellos casos donde las preferencias deben ser actualizadas frecuentemente y rápidamente.

- b) Técnicas de clustering o encasillamiento (Breese, Heckerman y Kadie, 1998): permiten realizar una clasificación utilizando patrones de grupos homogéneos. Es útil cuando el número de grupos es pequeño. Este tipo de técnica permite identificar grupos de usuarios que tienen preferencias similares para crear el grupo o cluster. Una vez creado los clusters, las predicciones pueden realizarse agregando las opiniones de otros usuarios pertenecientes a ese cluster. Algunos métodos más sofisticados tienen en cuenta la posibilidad de que un usuario pueda pertenecer parcialmente a varios clústeres, obteniendo la predicción mediante la agregación de los clústeres participantes y ponderándolo al grado de pertenencia a cada cluster.
- c) Horting (Wolf et al., 1999): interesante técnica basada en grafos en la cual los nodos son usuarios y las aristas entre nodos son indicadores de los grados de similitud entre varios usuarios. Su gran aportación es la utilización de técnicas de recorrido del grafo que exploran relaciones transitivas que los algoritmos de vecindad más cercana no tienen en cuenta.
- d) Basado en ítems (Sarwar et al., 2001): partiendo de los ítems que el usuario ha valorado, calcula valores de predicción que le daría el usuario a ítems similares, para así seleccionar aquellos más cercanos al modelo ya calculado. De esta forma, se obtiene la

valoración global con la media de las calificaciones del usuario en productos similares.

- e) Redes neuronales (Bueno, 2002): proporcionan una buena forma de representar el conocimiento en tareas de recuperación de información. Existen una gran cantidad de estudios de las redes neuronales aplicadas al filtrado colaborativos basado en el uso de perceptrones multicapa.
- f) Máquinas de soporte vectorial: cada ítem es mapeado en un espacio vectorial. Son métodos con bajo coste a nivel computacional.
- g) Similitud basada en el coseno: determina la similitud de dos vectores en un espacio multidimensional. El espacio puede describir características del usuario o de ítems. La similitud se mide calculando el ángulo entre dos vectores por medio del producto escalar.
- h) Correlación de Pearson: se trata de una métrica típica de similitud entre funciones de preferencia de usuarios o distancias de vectores.
- i) Basados en reglas de asociación: permite analizar patrones de preferencia entre los ítems, para recomendarlos a partir de otros ítems seleccionados. Las reglas de asociación expresan la relación de un determinado ítem, que es utilizado habitualmente, junto con otros ítems. Permiten formar una representación compacta de datos de preferencia que mejora la eficiencia. Suele utilizarse en aquellos casos donde los intereses no cambian rápidamente.

3.2.1.2. Basado en memoria.

Este tipo de sistemas de recomendación se basa en la colección completa de ítems valorados por el usuario con anterioridad a la solicitud de recomendación. Por tanto, el valor de predicción de interés de un ítem, no utilizado por un usuario concreto, se calcula como una agregación de las valoraciones de otros sobre dicho ítem (generalmente, se utiliza los k usuarios más parecidos al que solicita la recomendación).

Para la realización de estos cálculos se aplica el algoritmo de vecinos cercanos, estimándose una media ponderada de los vecinos más cercanos que hayan valorado ese producto. Este tipo de técnica presenta como ventaja la capacidad de utilizar rápidamente la información actual del sistema sobre las preferencias de los usuarios.

El principal inconveniente se produce cuando se incrementa el número de usuarios, ya que los recursos necesarios para realizar las recomendaciones también crecerán, aumentando de forma inaceptable el tiempo necesario para la realización de recomendaciones. Para disminuir este problema es habitual utilizar herramientas heurísticas que seleccionen vecinos óptimos dentro de una amplia población de usuarios. También se está comenzando a trabajar en técnicas difusas basadas en FCA-AFC⁴⁹ para sistemas de recomendación colaborativos en la búsqueda eficiente de vecinos (Leiva et al., 2013a).

Para entender el funcionamiento de las técnicas colaborativas basadas en memoria, vamos a desarrollar un sencillo ejemplo de un sistema de recomendación de museos.

⁴⁹ Formal concept analysis. Análisis formal de conceptos.

El sistema generará las recomendaciones encontrando las correlaciones entre los diferentes usuarios del sistema. Con esta metodología se obtienen fácilmente ítems de interés que no conoce el usuario que solicita la recomendación, pero que sí han sido valorados por otros afines a él, obteniendo una predicción de la valoración que el usuario daría a esos productos.

Vamos a poner un ejemplo sobre la visita de cinco turistas a cinco museos del centro histórico de Málaga. En la tabla 3.5 tenemos las valoraciones realizadas por estos cinco turistas. El símbolo '+' indica que al usuario le ha gustado el museo y el símbolo '-' que no le ha gustado⁵⁰.

	Alejandro	Isabel	Paloma	Carlos	Antonio
Picasso Málaga	-	+	+	-	+
Carmen Thyssen	+	-	-	+	+
Centro Arte Contemporáneo	-	? ?	+	-	+
Artes y costumbres populares	+	+	+	+	+
Catedralicio	+	-	-	+	-

Tabla 3.5. Valoraciones de turistas

Si queremos predecir si le interesaría a Isabel visitar el Centro de Arte Contemporáneo, deben buscarse usuarios que tengan un comportamiento similar al de Isabel. Si observamos la tabla, Paloma e Isabel tienen los mismos gustos, por lo que podemos deducir que a Isabel debe interesarle visitar el Centro de Arte Contemporáneo porque

⁵⁰ Para simplificar el problema se toma un valor binario para la valoración

le ha gustado a Paloma; pero esto sería una simplificación de la técnica que estamos describiendo.

La solución general debería encontrar el grado de correlación entre Isabel y el resto de usuarios, en vez de utilizar solo aquellos que tienen gustos más parecidos a los de Isabel, usando una media de las valoraciones para generar las recomendaciones. El peso dado a cada usuario coincidirá con el grado de correlación entre éste e Isabel. En general la valoración puede ser un valor a elegir entre una escala lingüística⁵¹ o un valor numérico, por ejemplo entre 0 y 10, no teniendo que ser tan radical asignando exclusivamente un valor binario (1 para '+', 0 para '-').

En el caso de valores numéricos es muy utilizado el coeficiente de correlación de Pearson. Así, denominamos $R_{i,j}$ a la valoración del usuario i sobre el ítem j , obteniendo la correlación entre el usuario u_1 y u_2 de la siguiente forma:

$$r(u_1, u_2) = \frac{\sum_{i \in I} (R_{u1,i} - \bar{R}_{u1}) \cdot (R_{u2,i} - \bar{R}_{u2})}{\sqrt{\sum_{i \in I} (R_{u1,i} - \bar{R}_{u1})^2 \cdot (R_{u2,i} - \bar{R}_{u2})^2}}$$

Donde \bar{R}_u es el valor medio de las valoraciones realizadas por el usuario u . Un algoritmo colaborativo predeciría que a Isabel le gustará el Centro de arte contemporáneo basándose en las recomendaciones del resto de turistas. También pueden utilizarse técnicas para encontrar correlaciones entre los ítems valorados. Así, el Centro de arte contemporáneo y el Museo Picasso Málaga tienen una correlación

⁵¹ Un ejemplo de escala lingüística podría ser Muy Buena, Buena, Normal, Regular y Mala.

perfecta, por lo que podríamos deducir que a Isabel le resultaría atractivo visitar el Centro de arte contemporáneo porque a ella le gustó el Museo Picasso. De forma análoga, se puede calcular la correlación entre los museos utilizando otra vez el coeficiente de Pearson y así realizar las predicciones, basándose en la media ponderada de las valoraciones de los museos.

3.2.2. Sistemas de recomendación basados en contenido

Esta técnica consiste en describir los ítems que se desean recomendar mediante atributos, propiedades, etc. Los sistemas de recomendación basados en contenido tratan de explotar esta información mediante la utilización de técnicas de recuperación y de filtrado.

Como ejemplo, podemos destacar los sistemas documentales que realizan recomendaciones basándose exclusivamente en el contenido de estos, es decir, cuando un usuario considera que un documento es interesante, se extraen del mismo un conjunto de palabras claves que describen e identifican el documento. Si la evaluación es positiva, se utilizarán estos parámetros para buscar documentos similares. Por tanto, este tipo de sistemas realiza las recomendaciones basándose únicamente y exclusivamente en la información que se tenga sobre el usuario y sus evaluaciones.

En la red podemos encontrar gran cantidad de información sobre un tema, pero no toda es relevante o interesante para las preferencias de un usuario determinado. Es cierto que existen diferentes mecanismos de orientar la búsqueda a las preferencias, pero también es cierto que los

resultados que devuelven son excesivos y muchos de ellos no son de interés para el usuario. El problema de encontrar información relevante dentro de una gran cantidad de datos se conoce con el nombre de sobrecarga de información. Esto requiere que aparezcan herramientas o técnicas para facilitar la labor de búsqueda a los usuarios.

Los sistemas de recomendación basados en contenido se utilizan fundamentalmente en la recuperación de datos en Internet, donde la información existente es muy grande y además creciente y cambiante. Una vez recuperada la información, debe realizarse una fase de filtrado, que consiste en considerar la relevancia de los elementos para el usuario.

Los sistemas de recomendación que utilizan esta técnica analizan las descripciones de los ítems que han sido valorados, para a partir de ellos predecir si un ítem puede adaptarse a sus preferencias. Cada ítem llevará asociado un conjunto de atributos que lo describe, así de una obra de arte podemos tener como atributos el autor, estilo, técnica, etc. Analizando las características que ha valorado positivamente un usuario, se construye un perfil que es utilizado para buscar ítems que puedan satisfacer sus preferencias.

Existen diferentes algoritmos basados en este tipo de técnica. Para mostrar su funcionamiento, nos centraremos en el algoritmo Winnow (Blum, Hellerstein y Littlestone, 1995; Pérez, 2008) a través de un ejemplo.

En la tabla 3.6 podemos ver las diferentes valoraciones de Isabel sobre los museos que ha visitado. Cada museo se ha estructurado en

función de las características del tipo de obras que posee y el objetivo es averiguar si un determinado museo debe recomendarse o no.

	Pintura abstracta (x ₁)	Pintura clásica (x ₂)	Escultura (x ₃)	Regional (x ₄)	Isabel
Picasso Málaga	S	S	S		-
Carmen Thyssen		S			+
Centro Arte Contemporáneo	S				-
Artes y costumbres populares		S	S	S	+
Catedralicio		S	S		?

Tabla 3.6. Valoraciones de un turista sobre museos

Cada atributo x_i (en el ejemplo pintura abstracta, pintura clásica, escultura y obras regionales), se trata como una variable booleana⁵² que tendrá un valor unitario, cuando esa característica esté incluida en la descripción del museo y cero en caso contrario, asignándole el algoritmo un peso w_i , a cada una de las características. Así, en la tabla 3.6 observamos que por ejemplo el museo Picasso contiene obras de pintura abstracta, clásica y esculturas, pero Isabel lo valoró negativamente.

Inicialmente para cada atributo se toman un peso de valor 1 y el algoritmo realiza un proceso de aprendizaje a partir de las visitas realizadas por el usuario, recalculando los pesos asociados a cada característica dependiendo de si la visita realizada fue de su agrado o no.

⁵² Se ha utilizado valoraciones binarias para simplificar el ejemplo

Finalizado el proceso de aprendizaje recomendará el ítem si cumple la siguiente inecuación:

$$\sum_{i=1}^n w_i \cdot x_i > umbral$$

Donde n es el total de características y el umbral, el valor a partir del cual se considera que debe recomendarse un determinado ítem.

El algoritmo descrito es el siguiente:

Para cada atributo i

$w_i = 1$

FinPara

Para cada visita realizada por el usuario u

Calcular $A = \sum_{i=1}^n w_i \cdot x_i$

Si $A > umbral$ y no le gustó la visita

entonces

Para cada atributo i

$w_i = w_i / 2$

FinPara

Finsi

Si $A < umbral$ y le gustó a visita

entonces

Para cada atributo i

$w_i = w_i \cdot 2$

FinPara

Finsi
FinPara

Una vez realizado este proceso de aprendizaje en el que se van ajustando los diferentes pesos de cada uno de los atributos, podemos conocer si al usuario hay que recomendarle o no un determinado producto. Para ello se suman los pesos asociados por las características que cumple un determinado ítem y si el valor supera el umbral se le recomienda y si no lo supera, no se le recomienda.

En el ejemplo de la tabla 3.6 se obtiene la suma de los pesos finales de los atributos pintura clásica y escultura, en caso de superar el umbral se le recomendará a Isabel la visita del Museo Catedraliceo.

Como conclusión a esta sección, podemos afirmar que los sistemas de recomendación basados en contenido tienen como principal inconveniente que acumulan bastante información sobre las preferencias de los usuarios que usan el sistema. Otro problema de esta técnica es el estar limitado por las características asociadas que describen a los ítems que puede recomendar.

Los sistemas de recomendación exclusivamente colaborativos no toman en consideración las características que describen el ítem, pero cuando este tipo de sistemas tienen en cuenta este factor, suelen ofrecer resultados más precisos (Bueno, 2002) que los sistemas basados exclusivamente en contenido.

Por otro lado, un sistema de recomendación basado exclusivamente en contenido al estar preparado para un usuario que utiliza un tipo

determinado de ítem, no le recomendará otras clases de ítems a no ser que tenga atributos compartidos.

3.2.3. Sistemas de recomendación demográficos

Los sistemas de recomendación demográficos se basan en conocer información personal de los usuarios que utilizan el sistema (edad, nivel cultural, sexo, etc.) y a partir de estos datos, realizar recomendaciones que hayan sido evaluadas positivamente por personas con características demográficas similares.

Este tipo de sistemas intenta asociar estereotipos a partir de la información demográfica y realiza recomendaciones acordes al perfil del usuario, relacionándolo con algunos de los estereotipos existentes. La organización de los estereotipos suele ser una estructura jerárquica, de manera que se permitan realizar razonamientos de inferencia (Pazzani, 1999).

Algunos métodos emplean metodologías complejas para la predicción de valoraciones de preferencia en las recomendaciones, destacando la técnica de descomposición en valores singulares de los vectores de usuario con información demográfica y la técnica basada en la utilización de la matriz de valoraciones (Vozalis y Margaritis, 2007). Con posterioridad se utiliza esta información para aplicar medidas de correlación no sólo de la información demográfica, sino también de las similitudes de las valoraciones.

Otros sistemas utilizan técnicas de hibridación que integran recomendaciones basada en contenido, filtrado colaborativo y

demográfico (Schiaffino y Amandi, 2009). En el sistema propuesto por Schiaffino y Amandi se consideraban las características demográficas del usuario y de los distintos servicios que se recomendaban, para asociar un valor de utilidad del producto al usuario.

Actualmente solo se aplican en dominios muy concretos, debido a las limitaciones que presenta su utilización. Los principales inconvenientes son:

- a) Dificultad de disponer de los datos demográficos necesarios.
- b) Algunos usuarios consideran un ataque a su intimidad y privacidad el proporcionar datos personales.

3.2.4. Sistemas de recomendación basados en conocimiento

Estos tipos de sistemas utilizan técnicas de razonamiento basados en casos para poder inferir las recomendaciones, aplicando mecanismos que permitan la obtención de analogías. Este tipo de técnica poseen la siguiente formalización (Hinkle y Toomey, 1994):

- a) Dado un determinado tipo de problema, el sistema debe recuperar aquellos casos similares a este y que han sido resueltos eficientemente con anterioridad.
- b) Se deben reutilizar las soluciones de problemas anteriores al nuevo.
- c) Debe probarse que la solución obtenida puede ser aplicada al problema en cuestión.
- d) Si la solución resulta satisfactoria, entonces debe almacenarse como un nuevo caso para resolver problemas futuros.

Los sistemas de recomendación basados en conocimiento deben tener información sobre los diferentes usuarios y los ítems que pueden recomendarse, para poder realizar las recomendaciones a partir del razonamiento de qué ítems satisfacen las preferencias del usuario. Para llevar a cabo este tipo de técnica es muy habitual realizar recomendaciones de ítems similares a los que le gustaban al usuario.

Este tipo de sistemas plantean algunas ventajas sobre los sistemas de recomendación clásicos descritos anteriormente. Así, los sistemas de recomendación colaborativos requieren una gran cantidad de valoraciones de los ítems para poder realizar las recomendaciones. Este problema también se presenta en los sistemas de recomendación que utilizan algoritmos de aprendizaje, que necesitan grandes cantidades de información. Así, para que el sistema pueda generar un modelo que represente las preferencias del usuario, se necesita una cantidad de información significativa tanto de acciones como de valoraciones anteriores. Los sistemas de recomendación basados en conocimiento no presentan este inconveniente, ya que aunque necesitan grandes cantidades de información, no están basadas en valoraciones de otros usuarios.

Por tanto, este tipo de sistemas son un importante complemento para otros sistemas híbridos de recomendación, sobre todo en aquellos casos con poca información sobre ítems utilizados anteriormente por el usuario. El principal inconveniente que presenta es que las recomendaciones siempre serán similares y el efecto sorpresa no se presentará en ningún caso.

Un caso especial de los sistemas de recomendación basados en conocimiento son los que utilizan el razonamiento basado en casos, que se fundamentan en tres etapas:

- a) Fase de recuperación de necesidades. El usuario especifica al sistema las características que deben tener los ítems en los que está interesado.
- b) Fase de búsqueda de ítems. El sistema debe buscar productos que tengan características similares a las descritas por el usuario en la fase anterior. El conjunto de ítems recuperados será el resultado de esta fase.
- c) Fase de refinamiento. Si los ítems recomendados no cumplen las expectativas del usuario, éste debe refinar aún más sus preferencias y se repetirá la fase primera. Esta técnica finaliza cuando el usuario encuentra los ítems que satisfacen sus preferencias.

Con estos sistemas se pueden obtener mejores resultados si se complementan con una mayor información del usuario. Esta podría venir de información demográfica o específica del propio usuario. Además, si la información de los ítems a recomendar está estructurada en características y atributos, esto redundará en la eficiencia de las recomendaciones generadas. Por ejemplo, si a un turista le gusta el deporte acuático y solicita una recomendación de un restaurante, podría ser recomendado alguno que esté en la costa.

Para poder tener en cuenta este tipo de información se requieren expertos que describan cada ítem de forma completa y estructurada;

dependiendo por tanto la calidad de la recomendación, de la calidad de este tipo de conocimiento aportado al sistema.

Otro elemento muy importante para que este tipo de sistema dé resultados óptimos, es disponer de herramientas con una interfaz sencilla y amigable, que permita al usuario de forma ágil y rápida proporcionar al sistema las consultas que solicita (Leiva et al., 2008) y de esta forma, se puedan realizar recomendaciones eficientes.

Es muy habitual que el sistema solicite al usuario ejemplos cercanos a las necesidades para poder realizar esa recomendación. Si el conjunto de ejemplos es numeroso, corremos el riesgo de que deje de utilizar el sistema por tener que dedicarle excesivo tiempo.

Otro problema de este tipo de técnica es que el usuario deje de utilizar el sistema al sentirse abrumado cuando tiene que describir muchas características de ítems. En la fase de refinamiento, quizá tenga que añadir nuevas características sobre los tipos de ítems que desea, haciendo más tediosa la tarea. Si por el contrario el número de atributos que se solicita al usuario es limitado, quizá le coarte las posibilidades.

A pesar de estos inconvenientes, los sistemas de recomendación basados en conocimiento han tenido bastante éxito cuando se solicitan ítems de un único dominio, como por ejemplo restaurantes, canciones, etc. (Burke, 2002). Aunque, como hemos dicho anteriormente, es muy habitual utilizarlo como complemento en sistemas de recomendación híbridos.

Por tanto, los sistemas de recomendación basados en conocimiento tienen como principales inconvenientes:

- a) Necesidad de adquirir el conocimiento
- b) Incapacidad para descubrir los grupos o tipos de usuarios.

Estos sistemas ofrecen buenos resultados en demandas puntuales (Pérez, 2008), pues requiere menos información del usuario que en otro tipo de técnicas. Un importante beneficio es que los tipos de recomendaciones que realiza dependerán de la base de conocimiento que tenga, siendo muy variado el tipo de recomendación que puede realizar.

3.3. Sistemas de recomendación aplicados al contexto

La importancia de la información contextual ha sido reconocida por investigadores y profesionales en muchas disciplinas como computación móvil, minería de datos, gestión, etc. La mayoría de los sistemas de recomendación actuales se centran en recomendar los temas más relevantes para los usuarios, sin tener en cuenta la información contextual. Algunos investigadores (Adomavicius et al., 2010) están comenzando a sostener que la información contextual relevante es muy importante en los sistemas de recomendación, siendo necesario tener en cuenta esta información en la prestación de recomendaciones.

Se consideran tres tipos de paradigmas de recomendación: prefiltrado contextual, postfiltrado contextual y modelado contextual (Adomavicius et al, 2010; Leiva et al., 2013a). Estos algoritmos incorporan la información contextual en el proceso de recomendación.

La mayoría de los sistemas de recomendación se centran en recopilar información relevante del usuario, para poder realizar una recomendación. Normalmente no se tienen en cuenta factores como lugar, hora, compañía, etc. Dicho de otra manera, tradicionalmente los sistemas de recomendación son aplicaciones que manejan dos tipos de entidades: usuarios y elementos a recomendar, no utilizando el contexto en la prestación de las recomendaciones.

Sin embargo, en muchas aplicaciones, como podría ser la recomendación de un paquete de vacaciones de un sistema web, puede que no sea suficiente con considerar al usuario y elementos a recomendar. Es muy importante incorporar información contextual en el proceso de recomendación para poder recomendar elementos que a los usuarios les interesan en determinadas circunstancias. Por ejemplo, utilizando el contexto temporal, un sistema de recomendación de viajes proporcionará una recomendación de vacaciones diferente en invierno que en verano.

Estos tipos de informaciones son extrapolables a otros campos como el marketing, donde se ha investigado el comportamiento de los consumidores. Se ha demostrado que la toma de decisiones, en lugar de ser invariable, es circunstancial al contexto (Adomavicius et al., 2010). Por tanto, una predicción necesita sin duda de las preferencias del consumidor, pero también depende del grado en que se haya incorporado la información relevante contextual al método de recomendación.

Más recientemente, algunas empresas empezaron a incorporar información contextual en sus motores de recomendación. Por ejemplo, al seleccionar una canción para el cliente, Sourcetone⁵³ toma en consideración el estado de ánimo del oyente (el contexto) que este mismo especifica. Sin embargo, todavía muchas de las soluciones de recomendación existentes en el mercado no ven claro si el contexto es importante para las aplicaciones de recomendación (Adomavicus et al., 2010).

Como conclusión, el contexto debe incluirse en los sistemas de recomendación, teniendo en cuenta la dependencia del dominio de aplicación y de los datos disponibles, ya que es indudable que cierta información contextual puede ser útil para proporcionar mejores recomendaciones (Baltrunas y Ricci, 2009; Adomavicius et al., 2010; Leiva et al., 2013a).

3.3.1. Definición de contexto

El contexto es un concepto multifacético que ha sido estudiado a través de diferentes investigaciones y disciplinas, incluyendo ciencias de la computación (principalmente en la inteligencia artificial), ciencia cognitiva, lingüística, psicología y ciencias organizativas. De hecho, la conferencia (CONTEXTO⁵⁴), se dedicó exclusivamente a estudiar este tema e incorporó a diversas ramas de la ciencia como medicina, derecho y las ciencias empresariales. Un investigador de negocios prestigioso como K. Prahalad declaró que "la capacidad para alcanzar a los clientes en cualquier lugar y en cualquier momento implica que las empresas

⁵³ www.sourcetone.com

⁵⁴ <http://context-07.ruc.dk>

deben ofrecer productos no sólo competitivos, sino también únicos y además en tiempo real, haciendo que la experiencia del cliente tenga en cuenta su contexto. Esto debe ser uno de los objetivos principales para los profesionales" (Prahalad, 2004).

El contexto ha sido estudiado en múltiples disciplinas y cada una de ellas tiende a tener su propia visión. Esto hace que posean matices las distintas definiciones que cada campo aporta respecto a la definición genérica de contexto ("Condiciones o circunstancias que afectan a alguna cosa").

Se han presentado y analizado más de 150 definiciones diferentes de contexto de los diferentes campos (Bazire y Brezillon, 2005), lo cual no es sorprendente, dada la complejidad y el carácter multifacético del concepto.

Dado que nuestro objetivo de estudio son los sistemas de recomendación y como el concepto general de contexto es muy amplio, a continuación lo trataremos en aquellos campos que están relacionados con los sistemas de recomendación, como son la minería de datos, personalización de comercio electrónico, sistemas móviles contextuales, bases de datos, marketing y negocios (Palmisano, Tuzhilin y Gorgoglione, 2008):

- a) Minería de datos. En este campo de investigación, el contexto se define comúnmente como los eventos que caracterizan las etapas de la vida de un cliente y que pueden determinar cambios en sus preferencias, estado y valoración (Berry y Linoff, 1997).

b) Personalización de comercio electrónico: el contexto se define como las circunstancias que provocan una intención de compra realizada por parte de un cliente en una aplicación de comercio electrónico. Diferentes intentos de compra pueden dar lugar a diversos tipos de comportamientos (Palmisano, Tuzhilin y Gorgoglione, 2008). Por ejemplo, un mismo cliente puede comprar el mismo producto teniendo en cuenta diferentes razones: un libro para mejorar sus conocimientos, como regalo, ocio, etc. Para hacer frente a las diferentes intenciones de compra, se construye un perfil diferente de cliente por cada contexto de compra, de esta forma estos perfiles separados se utilizan para la construcción de modelos de predicción, que son independientes del comportamiento del cliente en contextos y segmentos específicos. Dicha segmentación contextual de los clientes es muy útil, ya que resulta ser un mejor modelo de predicción a través de diferentes aplicaciones de comercio electrónico.

Los sistemas de recomendación también están relacionados con la personalización del comercio electrónico, ya que permite generar recomendaciones personalizadas de diversos productos y servicios a los clientes. La importancia de la inclusión y el uso de la información contextual en los sistemas de recomendación ha sido demostrado por Adomavicius, para ayudar a aumentar la calidad de las recomendaciones en ciertos entornos (Adomavicius et al, 2010).

- c) Sistemas móviles contextuales: inicialmente se definió como la ubicación del usuario, la identidad de las personas cercanas a él, los objetos a su alrededor y los cambios en estos elementos (Schilit y Theimer, 1994). Posteriormente se han añadido otros factores a esta definición, como por ejemplo la fecha, temporada y temperatura (Brown, Bovey y Chen, 1997), intereses físicos y conceptuales para un usuario (Ryan, Pascoe y Morse, 1997), además de su estado emocional (Dey, Abowd y Salber, 2001). Esta información contextual es crucial para proporcionar una amplia gama de servicios basados en la localización (LBE Location Based Services) a los clientes de telefonía móvil (Schiller y Voisard, 2004). Por ejemplo, un teatro puede recomendar entradas con un gran descuento media hora antes del inicio del espectáculo (ya que estas entradas no se venderían después del inicio del espectáculo) y enviar esta información a los smartphones de los visitantes cercanos. Por tanto, el tiempo y la ubicación constituyen información contextual en esta solicitud.
- d) Bases de datos: las capacidades contextuales en los sistemas de manejo de base de datos se definen mediante la incorporación de las preferencias del usuario, utilizando las diferentes respuestas a las consultas a las bases de datos en función del contexto en el que han sido expresadas (Stefanidis, Pitoura y Vassiliadis, 2007).
- e) Marketing y negocios: los investigadores de marketing han sostenido que el proceso de compra está condicionado al contexto en el que se realiza la transacción, ya que el mismo cliente puede adoptar diferentes estrategias de decisión y preferir diferentes

productos o marcas, dependiendo del contexto. Los consumidores varían sus reglas de toma de decisiones, debido a la utilización del producto y la situación de compra. Por lo tanto, la precisión en las predicciones de las preferencias del consumidor deben depender del grado en que hayamos incorporado la información contextual. En la literatura de Marketing, el contexto se define como una compleja tarea en la elección de la estrategia de marca (Bettman, Luce y Payne, 1991).

Una vez analizado el contexto en los diferentes ámbitos de conocimiento, podemos establecer una taxonomía, según la cual el contexto se puede clasificar desde dos puntos de vistas, según la representación y según la interacción (Dourish, 2004). Desde el punto de vista de *la representación*, el contexto se define como un conjunto predefinido de atributos observables y una estructura (o esquema de la base de datos, utilizando la terminología), la cual no cambia significativamente a lo largo del tiempo. En otras palabras, el punto de vista de *representación* supone que los atributos contextuales son identificables y conocidos *a priori* y por tanto, pueden ser utilizados en las aplicaciones basadas en contexto. En contraste, la vista *interacional* supone que el comportamiento del usuario es inducido por un contexto subyacente, que no es necesariamente observable. Además, se asume una relación *bidireccional* entre las actividades y contextos fundamentales: actividades que son influenciables por el contexto y también diferentes actividades que dan lugar a diferentes contextos.

3.3.2. El contexto en los sistemas de recomendación

El proceso clásico de recomendación intenta estimar la calificación por parte de un usuario de un elemento que no ha sido valorado. Una vez que se realiza esta estimación para todo el espacio usuario-elemento, el sistema puede recomendar el elemento de mayor valoración (o k mejores artículos) para cada usuario. A estos sistemas tradicionales se les denomina de dos dimensiones (2D), dado que solo consideran las dimensiones usuario e ítem en el proceso de recomendación (Adomavicius et al., 2010).

Los sistemas de recomendación contextuales (CARS) se ocupan de modelar y predecir los gustos y preferencias del usuario, mediante la incorporación de información contextual al sistema de recomendación como categorías adicionales de datos. Estas preferencias son presentadas no sólo en función de los usuarios y los ítems, sino también del contexto. En otras palabras, la predicción de interés de un ítem a un usuario se obtendrá mediante una función basada en tres variables (usuario, ítem y contexto).

Para ilustrar este concepto, podemos considerar como ejemplo un sistema de recomendación de películas, donde las dimensiones aplicables al sistema serían:

- Película: el conjunto de todas las películas que pueden ser recomendadas. Cada película que puede recomendarse está definida por los siguientes atributos: identificador, título, duración, año, director, actor principal, actriz principal y género.

- Usuario: las personas a las que se les va a recomendar películas. Cada usuario podría definirse mediante identificador, nombre, dirección, edad, género y profesión.
- Contexto: en el presente ejemplo supondremos que el contexto se describirá mediante los siguientes atributos:
 - Salas de Cine: se trata de los diferentes recintos donde se exhiben las películas. Estará definida cada sala por un identificador, nombre, dirección, capacidad, ciudad, estado, país.
 - Hora: la hora donde la película puede ser vista o ha sido vista. Estaría definida por ejemplo por fecha, día de la semana y lugar de la semana. El atributo día de la semana será Lunes, Martes..., mientras que lugar de la semana puede ser día laborable o fin de semana.
 - Compañía: representa a la persona o grupo de personas con las que puede verse una película. Podría ser definido por los siguientes valores: solo, amigos, pareja, familia, compañeros de trabajo y otros.

Por tanto, cuando un usuario realiza una calificación de una película, deberá tener en consideración dónde, cómo, cuándo y con quién ha visto la película. Así, el tipo de película a recomendar puede ser diferente si va a ir un fin de semana con la familia o un día de diario con la pareja.

Como se observa en este ejemplo, la información contextual puede ser de diferentes tipos, donde cada uno de ellos define un cierto aspecto del contexto como puede ser el tiempo, localización, compañía, propósito

de una compra, etc. Además, cada tipo contextual tiene una estructura que refleja la complejidad de la información contextual, que puede ser definida de diferentes formas, entre ellas una de las más populares es la estructura jerárquica, la cual permite representar mediante árboles la información contextual. Este tipo de estrategia es muy habitual en la mayoría de los sistemas de recomendación basados en contexto y sistemas de perfiles (Adomavicius et al, 2005; Palmisano, Tuzhilin y Gorgoglione, 2008).

Varios autores apuestan por definir la información contextual como un conjunto de dimensiones contextuales, donde cada dimensión contextual K es definida como un conjunto de q atributos, $K = (K^1, K^2, \dots, K^q)$ que tienen una estructura jerárquica que captura un tipo particular de contexto. Los valores tomados por los atributos K^q definen los niveles más finos (mayor granularidad), mientras que K^1 define el nivel más grueso (menor granularidad). Por ejemplo, en la figura 3.1 se presenta una jerarquía de cuatro niveles para el atributo k , que especifica la intención en una transacción de compra. Mientras la raíz (nivel más grueso) de la jerarquía para k define compras en todos los contextos posibles, el siguiente nivel es definido por el atributo K^1 {personal, regalo}. En el siguiente nivel, el valor “personal” de atributo K^1 se divide en un contexto aún más detallado: compra realizada relacionada con el trabajo u otro tipo. Similarmente, el valor “Regalo” para K^1 puede ser dividido en regalo para una pareja o amigo y regalo para parientes y otros. Así, sucesivamente los atributos son divididos en niveles de jerarquía.

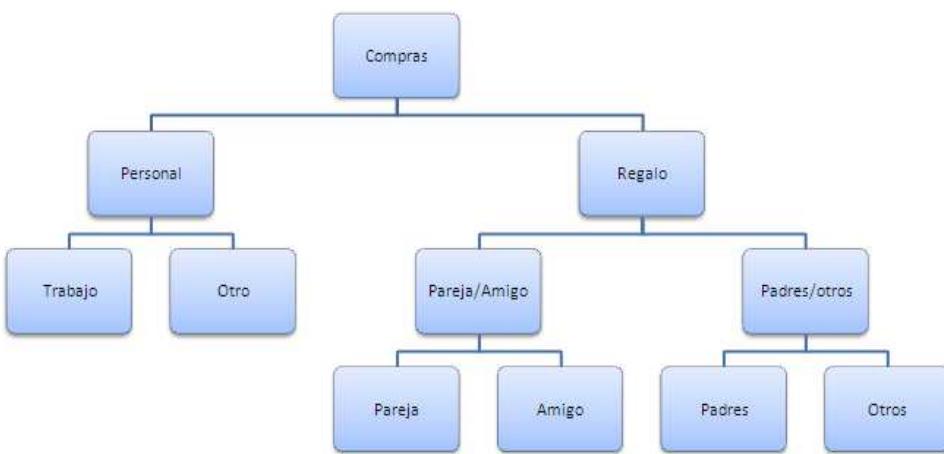


Figura 3.1. Estructura jerárquica de información contextual

La información contextual también puede ser definida utilizando otros tipos de mecanismos. Algunos autores plantean, además de la clásica dimensión de usuario e ítem, una dimensión contextual adicional, que es introducida aplicando un marco multidimensional (Adomavicius et al., 2010). Formalmente, D_1, D_2, \dots, D_n son dimensiones, de las cuales dos corresponden a usuarios e ítem, mientras que el resto será información contextual. Cada dimensión D_i es un subconjunto de un producto cartesiano de varios atributos (o campos).

Dada las dimensiones D_1, D_2, \dots, D_n , se define el espacio de recomendación como el producto cartesiano $S = D_1 \times D_2 \times \dots \times D_n$. Por otra parte, la calificación (rating) se representa como el conjunto ordenado de todos los posibles valores clasificados. Así, la función de clasificación es definida sobre el espacio $D_1 \times \dots \times D_n$ como:

$$R: D_1 \times \dots \times D_n \rightarrow \text{Calificación}$$

En el ejemplo, podemos definir una función de clasificación R en el espacio de recomendación (Usuarios, Items, Tiempo), especificando

como a un usuario $u \in \text{Usuarios}$ le gusta el ítem $i \in \text{Items}$ en el tiempo $t \in \text{Tiempo}$, $R(u,i,t)$.

Visualmente, las calificaciones $R(d_1, \dots, d_n)$ sobre el espacio de recomendación pueden ser almacenadas en un cubo multidimensional, como el que se muestra en la figura 3.2. Así, en el ejemplo la calificación $R(101,7,1)=6$ significa que para el usuario con Id 101 y el ítem Id 7, la calificación 6 es especificada para los días de diario.

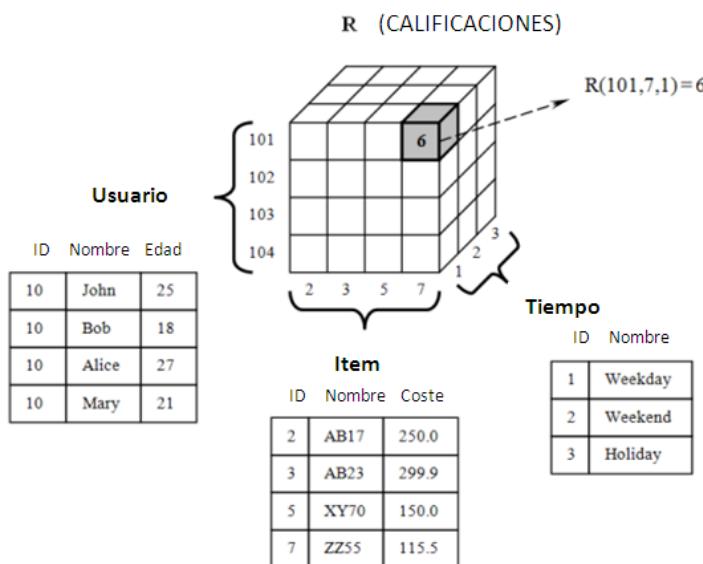


Figura 3.2. Modelo multidimensional

La principal diferencia de estos modelos es que mientras en la primera técnica descrita, la información contextual se define utilizando taxonomías jerárquicas, que pueden ser representadas como arboles (tanto balanceados como no balanceados), grafos acíclicos u otros tipos de taxonomías; la segunda técnica, para representar la información contextual, utiliza un modelo que es almacenado en cubos multidimensionales.

Otro aspecto a considerar en los sistemas de recomendación que utilizan información contextual consiste en los diferentes mecanismos para conseguir dichos datos, pudiendo obtenerse de diferentes maneras:

- a) Explícitamente: accediendo directamente a personas relevantes y otras fuentes de información contextual o recopilando esta información mediante cuestionarios u otros tipos de medios. Por ejemplo, un sitio web puede obtener información contextual, preguntando a personas mediante la cumplimentación de un formulario web o planteando algunas cuestiones específicas, después de proporcionar el acceso a determinadas páginas web.
- b) Implícitamente: no se debe hacer nada en estos casos en términos de interacción con el usuario u otras fuentes de información contextual. Así, la fuente de información contextual implícita debe ser accesible directamente y los datos se extraerán de ella. Un ejemplo podría ser obtener información a partir de datos del entorno, tales como el cambio de localización del usuario que es detectado por la compañía de teléfono móvil. Alternativamente, la información contextual temporal puede ser obtenida implícitamente a partir de fecha y hora de la realización de la transacción.
- c) Infiriendo o deduciendo el contexto: para ello se utilizan métodos estadísticos o de minería de datos. Por ejemplo, la identidad de una persona que cambia de canal de televisión (marido, esposa, hijo, etc.) puede no ser conocida explícitamente para una compañía de televisión, pero esto debería ser deducido con una precisión razonable al observar los programas de televisión vistos

y los canales visitados, utilizando diferentes métodos de minería de datos. Para deducir esta información contextual, será necesario construir un modelo predictivo (por ejemplo, un clasificador) y formarlos en los datos apropiados. El éxito de esta deducción de la información contextual depende muy significativamente de la calidad de dichos clasificadores, aunque esto puede variar significativamente según las diferentes aplicaciones (Palmisano, Tuzhilin y Gorgoglione, 2008, Noguera et al., 2012).

3.3.3. Paradigmas para incorporar el contexto en los sistemas de recomendación

El uso de la información contextual en los sistemas de recomendación se remonta a trabajos, en los que se partía de la hipótesis de que incluir conocimientos sobre tareas del usuario en el algoritmo de recomendación puede mejorar los resultados (Herlocker y Konstan, 2001). Así, si queremos recomendar libros como regalos para niños, entonces puede que deseemos especificar varios títulos que el niño ya tiene y suministrar esta información (por ejemplo mediante una tarea de perfiles) al sistema para calcular nuevas recomendaciones. Esta aproximación opera sobre el uso tradicional de dos dimensiones (espacios usuario e ítem), sin utilizar dimensiones contextuales. Sin embargo nos sirve como una ilustración del éxito que puede proporcionar información adicional, incorporada en un paradigma de filtrado colaborativo estándar.

Las diferentes técnicas para utilizar la información contextual en procesos de recomendación pueden ser categorizadas en dos grupos:

- a) Recomendación a través de consultas y búsquedas contextuales.
- b) Recomendación a través de preferencias y estimaciones contextuales.

El contexto mediante consulta y búsqueda ha sido utilizado por una gran cantidad de sistemas de recomendación turísticos y de móviles (Cena et al., 2006). Los sistemas que emplean este tipo de enfoque suelen usar la información contextual obtenida directamente del usuario, como por ejemplo mediante la especificación actual de estado de ánimo o bien por información del entorno (como puede ser hora, clima o ubicación), para poder consultar o buscar recursos (por ejemplo, restaurantes) y posteriormente presentar los mejores productos al usuario (por ejemplo, los restaurantes de la zona que están abiertos). Uno de los primeros ejemplos de este enfoque es el proyecto Cyberguide (Abowd et al., 1997), que desarrolló varios prototipos de guías turísticas para diferentes plataformas portátiles. Algunos otros ejemplos de guías turísticas basadas en contexto son: GUIDE (Breese, Heckerman y Kadie, 1998), INTRIGUE (Ardissono et al., 2003), COMPASS (Van Setten, Pokraev y Koolwaaij, 2004) y MyMap (De Carolis et al., 2009).

El segundo enfoque que utiliza la información contextual en el proceso de recomendación, es el que se realiza a través de las preferencias y las estimaciones, el cual representa una tendencia más actual en la literatura de sistemas de recomendación basados en el contexto (Adomavicius et al., 2005; Oku et al., 2006; Panniello et al., 2009). En contraste con el método anterior, éste intenta modelar y aprender las preferencias del usuario, observando la interacción de este

y otros con el sistema u obteniendo las preferencias mediante retroalimentación a partir del uso de varias recomendaciones previas. Para modelar las preferencias del contexto del usuario y generar recomendaciones, estas técnicas adoptan filtros típicamente colaborativos basados en el contexto o métodos de recomendación híbrida, donde aplican varias técnicas de análisis de datos, a partir de información obtenida mediante aprendizaje o minería de datos (como clasificadores bayesianos).

Es cierto que muchas aplicaciones combinan ambas técnicas o enfoques. Por ejemplo, UbiquiTO (Cena et al., 2006) es un sistema que representa una guía turística móvil, que suministra adaptación inteligente basada en información específica de contexto y que además utiliza técnicas basadas en reglas y modelos difusos, para adaptar el contenido de la aplicación basándose en las preferencias del usuario e intereses. Del mismo modo, News@hand (Cantador y Castells, 2009) utiliza tecnología semántica, para suministrar recomendaciones personalizadas de noticias, que se basan en consultas del usuario o cálculos de acuerdo al perfil del usuario o del grupo al que pertenece.

Normalmente en todos los sistemas de recomendación se utiliza una función de dos dimensiones basada en usuario e ítem, la cual toma datos de las preferencias como entrada y produce una lista de recomendaciones para cada usuario como una salida. La figura 3.3 presenta un resumen tradicional de un proceso de recomendación 2D, que incluye tres componentes: datos (entrada), sistema de recomendación 2D (función) y lista de recomendación (salida). Los

datos vienen determinados por las dimensiones usuarios (U), ítems (I) y calificaciones (R).



Figura 3.3. Componentes en un proceso de recomendación general

Los sistemas de recomendación tradicionales son construidos basándose en el conocimiento de las preferencias parciales de los usuarios, por ejemplo preferencias sobre algún conjunto de ítems (frecuentemente limitado). Además, los datos de entrada para el sistema están típicamente basados en registros de la forma <usuario, ítem, calificación>. En contraste, los sistemas basados en contexto son construidos a partir del conocimiento de las preferencias parciales contextuales del usuario (C), que tienen típicamente registros de la forma <usuario, ítem, contexto, calificación>, donde cada registro incluye no sólo como a un usuario le gusta un determinado ítem, sino también la información contextual en la cual el ítem fue consumido.

Por otra parte, además de la información descriptiva de los usuarios (datos demográficos), ítems (característica de los mismos) y calificaciones (información de calificación multicriterio), los sistemas de recomendación basados en el contexto también pueden hacer uso de diferentes tipos de atributos adicionales de contexto. Partiendo de la presencia de estos datos contextuales adicionales, surgen varias preguntas importantes: ¿Cómo debe la información contextual reflejarse a la hora de modelar las preferencias del usuario? ¿Puede ser reutilizada

la riqueza del conocimiento en sistemas de recomendación tradicionales (no contextuales) para generar recomendaciones basadas en el contexto?

Adomavicius presentó una clasificación de paradigmas de recomendación contextual distinguiendo tres formas, según el momento en el cual el contexto es utilizado por el sistema (Adomaviciu et al, 2010).

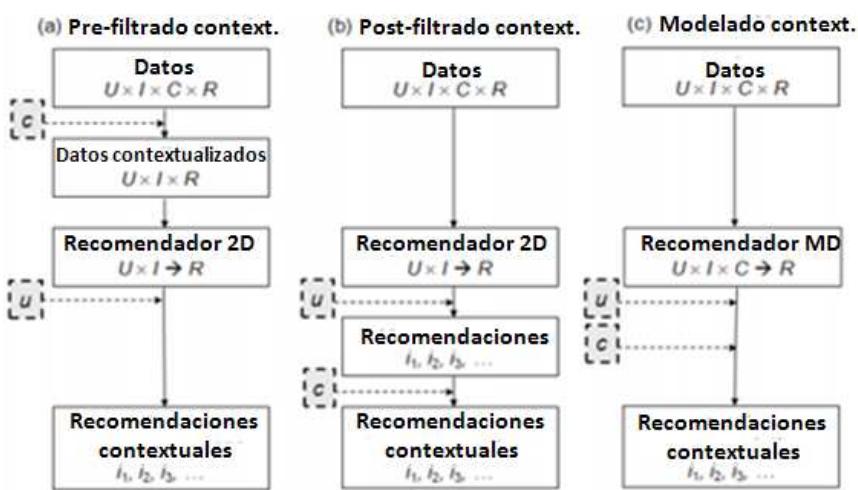


Figura 3.4. Paradigmas contextuales en SR

- Pre-filtrado contextual (o contextualización de la entrada de recomendación). En este paradigma de recomendación presentado en la figura 3.4a, la información contextual conlleva selección de datos para un contexto específico. En otras palabras, la información acerca del contexto actual c es utilizada para seleccionar o construir el conjunto de registros de datos. Una de las ventajas de emplear este modelo es que en los casos donde el número de ítems es grande, la utilización inicial del contexto

hará reducir drásticamente los ítems que participan en el proceso de recomendación posterior (Leiva et al., 2013a)

- b) Post-filtrado conceptual (o la contextualización de la salida de recomendación). En este paradigma de recomendación (presentado en la figura 3.4b), la información contextual es inicialmente ignorada y las calificaciones son predicciones usando cualquier sistema tradicional de recomendación 2D sobre los datos introducidos. Posteriormente, el conjunto resultante de recomendaciones es ajustado (contextualizado) para cada usuario utilizando la información contextual.
- c) Modelado contextual (o contextualización de la función de recomendación). En este paradigma de recomendación (presentado en la figura 3.4c), la información contextual se usa directamente en la técnica de modelación como parte de la estimación de las calificaciones.

3.4. Técnicas de hibridación

Todos los sistemas de recomendación poseen limitaciones, por lo que la técnica de hibridación es una solución para mejorar la eficiencia de las recomendaciones. Normalmente los sistemas híbridos mezclan dos o más técnicas de recomendación, para conseguir perfeccionar la precisión de este tipo de métodos (Leiva, Guevara y Rossi, 2012). Lo más habitual es combinar la técnica colaborativa con otras para evitar el problema de incremento (ramp-up). Las técnicas de hibridación más conocidas son:

- a) Uso de pesos: combina las recomendaciones generadas por cada sistema de recomendación, asignando a cada uno de ellos un peso dependiendo del sistema que las genera (Claypool et al., 1999). Algunos sistemas tienen pesos fijos y otros pueden cambiar el valor de éstos según las valoraciones que realicen los usuarios sobre las recomendaciones. La principal ventaja de este tipo de técnica es que explota todas las capacidades del sistema de forma sencilla. El inconveniente reside en considerar a todas las técnicas una importancia uniforme, lo cual no siempre tiene que ser cierto, pues en los casos donde un ítem tiene pocas valoraciones, los sistemas de recomendación basados en técnicas colaborativas no ofrecen buenos resultados. Un ejemplo es el sistema P-Tango (Claypool et al., 1999).
- b) Uso de técnicas “switch”: este tipo de modelo de hibridación alterna diferentes motores o técnicas dependiendo de la situación y el momento en el que se realiza la solicitud de la recomendación. La principal complejidad e inconveniente que posee es la determinación del criterio por el cual se utiliza un motor de recomendación u otro (Billsus y Pazzani, 2000). Una elección adecuada de estos criterios conllevará evitar las debilidades de unos motores y potenciar los beneficios de los otros. Un ejemplo es el sistema Daily Leaner (Billsus y Pazzani, 2000).
- c) Mezclas: este tipo de técnica combina los resultados obtenidos por los diferentes sistemas de recomendación implicados. Un ejemplo es el sistema PTV (Smyth y Cotter, 2000), que utiliza

técnicas colaborativas (para aprovechar información de preferencias de otros usuarios) y técnicas basadas en contenido (utilizando la descripción de los ítems a recomendar). El principal inconveniente es la necesidad de disponer de datos completos sobre las preferencias del usuario.

- d) Cascada: consiste en que el primer sistema obtiene unas recomendaciones que luego son refinadas por los siguientes sistemas. De esta forma se combinan las técnicas mediante procesos de filtrado de los resultados obtenidos por sistemas de recomendación aplicados anteriormente (Burke, 2002; Pérez, 2008). Una de las principales ventajas de este tipo de técnica es la eficiencia computacional, ya que excepto el primer sistema, el resto sólo tiene que recomendar entre los ítems seleccionados por el sistema anterior y no con el conjunto completo de ítems.

El uso de técnicas de hibridación permite por tanto, mejorar los resultados independientemente del método utilizado para realizar la hibridación (Burke, 2002). Es cierto que el problema de hibridar técnicas colaborativas y técnicas basadas en contenido conlleva la necesidad de convivir con el problema de incremento, pero es habitual utilizar este método de hibridación, ya que presentan mejores resultados que aplicarlas de forma independiente (Burke, 2002; Pérez, 2008).

Otra estrategia interesante consiste en utilizar modelos basados en conocimiento y utilidad en la hibridación puede ser una buena estrategia, ya que no presentan los problemas de incorporar nuevos usuarios o nuevos ítems (Pérez, 2008).

3.5. Recomendaciones para grupos

Como hemos comentado con anterioridad, la mayoría de los sistemas de recomendación se ocupan tradicionalmente de hacer recomendaciones de ítems a usuarios individuales. En estos últimos años se está empezando a trabajar en el desarrollo de técnicas que permitan proponer recomendaciones a grupos de usuarios simultáneamente (Jameson y Smyth, 2007; Leiva et al., 2012). Este tipo de sistema plantea como reto principal la necesidad de adquirir las preferencias del grupo, ayudando a la toma de decisiones y explicando el motivo o razones de esa elección. La mayoría de sistemas de recomendación de grupos tienen métodos de adquisición de información distintos a los que se aplican a los sistemas individuales, básicamente pueden dividirse en:

- a) Adquisición de preferencias sin especificación explícita: el sistema funciona con información adquirida implícitamente sobre los usuarios. Un ejemplo es el sistema Let's Browse, que recomienda páginas web a un grupo de usuarios que están navegando por Internet juntos (Lieberman, Dyke y Vivacqua, 1999).
- b) Especificación de preferencias explícitas: otros tipos de recomendadores sí necesitan especificación explícita de las preferencias de los usuarios. Un ejemplo interesante es el sistema propuesto por Jameson (Jameson y Smyth, 2007), que consiste en ayudar a un grupo de usuarios a acordar unas determinadas características de unas vacaciones que proyectan juntos. Otros sistemas plantean la puntuación de forma individual y cuando se realizan las recomendaciones se intenta minimizar el malestar,

basándose en la hipótesis de que la felicidad del grupo será igual a la del menos satisfecho de sus miembros (Shafer et al, 2007).

Cuando el sistema adquiere el conocimiento sobre los diferentes usuarios, puede adoptarse la especificación de las preferencias de la recomendación para el grupo. Entre los diferentes ejemplos podemos encontrar:

- Adaptive Radio (Chao, Balthrop y Forrest, 2005): se centra en las preferencias negativas. Pretende evitar que los ítems seleccionados sean especialmente negativos a los gustos de cualquier miembro del grupo.
- CATS (McCarthy et al, 2006): sistemas que comparten información sobre las preferencias. Así, cada miembro del grupo pueda conocer las preferencias de los otros miembros, con objeto de ahorrar tiempo para especificar sus propias preferencias, permitiendo llegar a un consenso y prever comportamientos de los otros componentes del grupo. Quizá este tipo de sistema tiene el inconveniente de que si una persona no quiere que salga un determinado tipo de ítem, puede calificarlo con la nota más negativa posible, consiguiendo que nunca pueda salir recomendado ese ítem, es decir, se puede manipular el sistema.

Por otra parte, dependiendo de diversas características del grupo, como pueden ser el tamaño y la homogeneidad, puede resultar muy complejo encontrar una recomendación adecuada para cada miembro del grupo (Jameson y Smyth, 2007). El caso más habitual es que el sistema escoja aquella opción que satisfaga al mayor número de usuarios de acuerdo a sus preferencias individuales. Por tanto, se hace necesario

encontrar métodos de agregación que permitan combinar la información sobre las distintas preferencias individuales de los usuarios, de forma que el sistema pueda obtener la recomendación idónea. Para resolver este problema, podemos encontrar diferentes métodos:

- a) Método simple: mezclar las recomendaciones individuales en una única lista (O'connor et al, 2001).
- b) Método de agregación: une las valoraciones para cada usuario y para cada ítem candidato. El sistema predice como cada componente del grupo evaluaría el ítem y devuelve una lista de ítems que tengan las valoraciones previstas más altas (McCarthy, 2002). Para ello debe tenerse en cuenta como objetivo maximizar la satisfacción media, penalizando aquellos que con media alta tenga algunas calificaciones bajas por parte de alguno de los usuarios del grupo.
- c) Método de construcción de perfil de grupo: consiste en la creación de un perfil especial para el grupo a partir de las características de todos los miembros del grupo (Leiva et al., 2012).

Las técnicas existentes en sistemas de recomendación para grupos se basan fundamentalmente en la utilización de funciones de agregación, es decir, acumular los intereses individuales de los diferentes miembros del grupo. Sin embargo, este tipo de estrategia ha sido criticada por varios autores, ya que no tienen en cuenta las relaciones de los diferentes miembros del grupo con el resto.

Añadir las interacciones entre los diferentes miembros del grupo, supone tener en cuenta información sobre el comportamiento de cada uno de ellos con el resto de miembros.

En general podemos utilizar dos tipos de técnicas para mejorar las recomendaciones para grupos: recomendaciones basadas en la personalidad y recomendaciones basadas en factores sociales.

- a) Recomendaciones basadas en la personalidad: la mayoría de las recomendaciones para grupos le dan la misma importancia a las preferencias de cada uno de los miembros. Esto no siempre es así, ya que no todas las personas de un grupo tienen la misma flexibilidad a la hora de aceptar una recomendación. Resulta muy interesante aplicar técnicas que tengan en cuenta la personalidad de los diferentes miembros del grupo, en concreto, la reacción que toma un individuo cuando el sistema le recomienda un tipo de producto diferente al que esperaba recibir (McCarty, 2002). Es importante destacar que no todos los grupos son iguales, no sólo respecto al número, sino también a que a veces un grupo está formado por personas con preferencias similares o antagónicas. En esta última situación, donde existen miembros del grupo incompatibles, debe resolverse la dificultad teniendo en cuenta rasgos de la personalidad del grupo, atendiendo a satisfacer al mayor número de usuarios y no perjudicando a ninguno de ellos. La satisfacción de un grupo, por tanto, no es la suma de las satisfacciones individuales. Es mejor recomendar una actividad que tenga una satisfacción media aceptable para todos, aunque pueda tener una suma de satisfacción inferior a otras actividades. Para tener en cuenta rasgos de personalidad o flexibilidad es muy común utilizar cuestionarios para caracterizar a los usuarios y de esta forma medir el comportamiento de cada uno de los

individuos en situaciones conflictivas, midiendo características relacionadas con el autoritarismo y el cooperacionismo. (Thomas y Kilmann, 1974). Así, el autoritarismo penaliza negativamente las diferencias entre las selecciones preferidas de otros miembros, mientras que el cooperacionismo recompensa las diferencias con las selecciones preferidas por otros componentes del grupo.

- b) Recomendaciones basadas en factores sociales: en este tipo de técnica se tienen en cuenta informaciones respecto del grado de confianza que tiene un usuario con el resto. Para ello, el uso de las redes sociales ofrecería una gran cantidad de información muy interesante (Gilbert y Karahalios, 2009; Younghoom y Kyuseok, 2014). Por ejemplo, si un usuario viaja con otro y a este último le gusta un determinado producto que no termina de satisfacer al primero, el hecho de que haya sido visitado por amigos comunes que lo han valorado positivamente, puede hacer que este producto tenga un valor de recomendación mayor que si los amigos comunes lo hubieran valorado negativamente. Existen muchos factores que pueden servir para el cálculo de la confianza: número de amigos en común, intensidad de la relación entre amigos, distancia social, número de fotos en las que aparece juntos, etc.

3.6. Aplicación de técnicas FCA-AFC⁵⁵ a sistemas de recomendación

Actualmente están empezando a aparecer técnicas de recomendación que utilizan herramientas basadas en el análisis formal de conceptos (Xiaohui y Murata, 2010; Nenova et al, 2013). La utilización de esta técnica puede ser usada para diversos objetivos, como por ejemplo para el cálculo de vecinos en un sistema de recomendación colaborativo (Nenova et al., 2013) o para filtrar un conjunto de productos que no cumplen unas reglas contextuales (Leiva et al., 2013a). En la presente sección vamos a repasar algunos de los fundamentos del análisis formal de conceptos (Wille, 1989; Ganter y Willie, 1999).

Se trata de un término que fue introducido por Rudolf Wille en 1982 y es una técnica de aprendizaje, basada en la extracción de estructuras conceptuales en un conjunto de datos. Se fundamenta en que todo concepto está conformado por:

- a) Extensión: todos los objetos que pertenecen a dicho concepto.
- b) Intención: todos los atributos compartidos por dichos objetos.

El marco en el que se establecen los conceptos se conoce como contexto formal. Consta de un conjunto de objetos, un conjunto de atributos o propiedades y una relación que informa sobre los atributos que posee cada objeto. El conjunto de los conceptos de un contexto formal tiene una estructura de retículo, lo que admite la posibilidad de representarlos en forma de jerarquía conceptual, que permitirá un análisis de las estructuras que revelará dependencias entre los datos.

⁵⁵ Formal concept analysis. Análisis formal de conceptos.

La utilización de estos tipos de técnicas está ofreciendo unos resultados óptimos en tareas relacionadas con análisis de datos, recuperación de información y descubrimiento de conocimiento en bases de datos. Además, FCA también está siendo utilizado en campos diversos como la deducción automática y sistemas de recomendación (Leiva et al., 2013a). Es habitual utilizar la sinergia del uso del conocimiento, partiendo de un repositorio que incluya el comportamiento de los usuarios y las propiedades de los ítems. Este modelo permite definir el perfil del usuario y del ítem, construyendo el conjunto de recomendación de candidatos utilizando técnicas FCA e inferencia extendida (Xiaohui y Murata, 2010).

3.7. Conclusiones

La actividad turística está íntimamente relacionada a las preferencias, intereses y gustos personales. Este es el motivo principal por el que muchas aplicaciones webs relacionadas con el turismo incorporan sistemas de recomendación. Es una forma de simular la interacción de un agente turístico con el turista o potencial cliente (Balabanovic y Shoham, 1997).

Existe una amplia bibliografía que defiende la necesidad de utilizar los sistemas de recomendación en servicios orientados al turismo (Billsus y Pazzani, 2000; Fesenmaier et al., 2003).

Quizás son menos abundantes las técnicas y servicios existentes para la creación de rutas turísticas, a partir de una serie de puntos recomendados dentro de una zona en la que se encuentra (Castillo et al.,

2008; Rodríguez, Molina y Caballero, 2008). Realmente la posibilidad de utilización de tecnologías móviles basadas en realidad aumentada que permitieran la creación de rutas según la posición en la que se encuentre el turista, sería un importante valor añadido a los servicios turísticos de una zona. Un proyecto similar al propuesto se creó (Deep Map, 2001) para desarrollar guías turísticas para webs y móviles sin utilizar realidad aumentada. Dentro de este proyecto se enmarca una aplicación (WebGuide, 2001) que permite obtener itinerarios para realizar visitas guiadas en la ciudad de Heidelberg, teniendo en cuenta las preferencias del turista. El sistema se encarga de identificar puntos de interés y calcula la ruta basándose en información geográfica, puntos de interés, medios de transporte seleccionados, así como restricciones que imponga el turista.

Tradicionalmente, la utilización de técnicas de recomendación a servicios turísticos conlleva ciertas dificultades, entre las que podemos destacar:

- a) La recomendación colaborativa es difícilmente aplicable a la mayor parte de esos sistemas. Para que se aplique correctamente, deberían los usuarios puntuar muchos elementos. Siendo tan grande la cantidad de puntos turísticos de interés en un destino, no parece nada sencilla la tarea de que el usuario puntúe suficientes elementos. La forma elegida para resolver este inconveniente está en que el turista valora tipos de puntos turísticos en su perfil y según va realizando visitas, las valoraciones que realice se irán añadiendo a su perfil. Además, las

calificaciones realizadas por los turistas redundarán en la valoración de los puntos para los diferentes estereotipos de turistas existentes. El punto fundamental está en clarificar adecuadamente cada turista qué perfil tiene asignado.

- b) La recomendación para grupos es muy importante. Habitualmente se viaja en grupo, por lo que sería ideal tener en cuenta las preferencias de todos los usuarios de dicho grupo. La creación de perfiles por grupos podría ser una opción que asociaría a un grupo el perfil de un usuario virtual, es decir, como si fuera de un individuo único.
- c) La recomendación ofrecida al usuario no depende únicamente de las preferencias de otros usuarios y del turista individual, sino que debe tener en cuenta la información del entorno. La recomendación no sólo debe basarse en intereses y preferencias, los atributos contextuales son de vital importancia para obtener una recomendación adecuada. Uno de los atributos contextuales intrínsecos de los sistemas de realidad aumentada será la detección de la posición en la que se encuentra el turista, para sugerir los puntos turísticos de interés que están en la zona o a una distancia mínima de donde se encuentre el turista.

Capítulo 4

Organización del conocimiento

4.1. Representación del conocimiento

La representación del conocimiento es una rama de la inteligencia artificial que se ocupa del diseño e implementación de sistemas para poder expresar el conocimiento en un dominio determinado. El mecanismo para conseguirlo se centra en la utilización de descripciones semánticas, con el objetivo principal de utilizar la información de estas descripciones en diferentes procesos computacionales de razonamiento de forma eficiente. La forma de representar el conocimiento viene determinado por el tiempo de razonamiento que se desee llevar a cabo (Blanco et al., 2006).

A continuación presentaremos secciones, donde definiremos conceptos relacionados con la representación del conocimiento.

4.1.1. Redes semánticas

Se trata de grafos acíclicos dirigidos en los que los conceptos se representan mediante nodos y las relaciones existentes entre ellos con enlaces o conectores. Inicialmente las redes semánticas se estructuraban y organizaban heterogéneamente dependiendo del tipo de sistema, para ello utilizaban un método de razonamiento basado en implementaciones específicas propietarias, sin utilizar lenguajes formales.

A principio de los años 80, se comienza a trabajar en la formalización de las redes semánticas mediante la creación de un lenguaje de representación del conocimiento basado en técnicas lógicas. Uno de los primeros lenguajes es KL-ONE, considerado la primera conceptualización formal de una red semántica (Brachman y Schmolze, 1985).

4.1.2. Representación de conocimiento basado en marcos

Los marcos son estructuras de datos que permiten representar conceptos y relaciones entre los mismos. Utilizan una terminología basada en atributos que permite identificar las relaciones establecidas en un dominio de aplicación específico (Minsky, 1975).

El conocimiento de un dominio se representa formalmente mediante entidades, que son descritas en clases e instancias. En un lenguaje basado en marcos, las clases y las instancias tienen asociadas un conjunto de atributos referidos al dominio concreto de una ontología. Sus principales características son:

- a) Organización jerárquica de clases.

- b) Una clase puede tener asociadas varias superclases.
- c) Las instancias heredan las propiedades y atributos de las clases a las que pertenecen.
- d) Posibilidad de declarar valores por defecto para dichos atributos.

4.1.3. Lógicas descriptivas

Las redes semánticas evolucionaron hacia los sistemas basados en marcos y posteriormente surgieron las lógicas descriptivas, las cuales utilizaban lenguajes que además de soportar la representación basada en marcos, permitía la formalización lógica mediante estructuras denominadas descripciones.

En el dominio de las lógicas descriptivas, una descripción es una expresión en un lenguaje formal que define un conjunto de instancias. Por tanto, una lógica descriptiva es un lenguaje que permite representar el conocimiento mediante una sintaxis y semántica, de forma que la sintaxis se utiliza para construir las descripciones, mientras que la semántica define el significado de cada una de las descripciones (Blanco, 2007).

Las lógicas descriptivas limitan el tipo de conocimiento que pueden formalizar, de forma que los problemas de decidibilidad son obviados en la implementación del sistema, además, cuando se formula una consulta a un sistema que ha sido formalizado mediante lógicas descriptivas, es posible que no se presenten resultados debido a que la búsqueda se realiza sobre un espacio infinito, es decir, pueden establecerse relaciones no decidibles.

Entre las principales funciones de las lógicas descriptivas están:

- a) Admitir formalismos lógicos para representar y expresar el conocimiento.
- b) Proporcionar métodos de razonamiento correctos y resolubles fundamentados en bases lógicas y teóricas sólidas.

Los lenguajes de representación de conocimiento más extendidos en el estado del arte son LOOM, KRIS y RACER (Haarlsev y Möller, 2001; Blanco, 2007), los cuales deben combinar la expresividad al formalizar el conocimiento para describir la semántica de un dominio concreto; además de proporcionar métodos de razonamiento resolubles para asegurar procesos de inferencia que descubran nuevos conocimientos.

4.2. Organización del conocimiento

En la presente sección vamos a definir las siguientes estructuras: taxonomías, mapas temáticos, tesauros y ontologías.

4.2.1. Taxonomías

Son estructuras que clasifican la información de forma jerárquica. La información se representa mediante clases, que se organizan en un árbol a partir de una única clase raíz. La estructura está construida de forma que al descender por el árbol de la taxonomía desde la raíz, se va accediendo a entidades más específicas; mientras que desde una entidad hoja hacia la raíz, se va accediendo hacia entidades o clases más generales (Montaner, López y De la Rosa, 2003).

Las taxonomías modelan relaciones de herencia, inclusión o pertenencia, que identifican conceptos basados en la teoría de grafos,

como profundidad, clases hojas y antecesores comunes más cercanos. Muchas de las capacidades de las taxonomías han sido aplicadas a los mapas temáticos.

4.2.2. Mapas temáticos

Los mapas temáticos sirven para organizar la información y facilitar su recuperación, permitiendo representar el conocimiento de forma normalizada (Le Grand, Soto y Dodds, 2001), modelar y representar el conocimiento de forma intercambiable y procesable, ofreciendo un marco unificado para la gestión del conocimiento y de la información. Además, posibilita la descripción de estructuras de conocimiento y las asocia a otros recursos de información.

Un mapa temático puede estructurarse como una colección de temas enlazados entre sí que definen una estructura denominada TAO formada por temas, asociaciones y ocurrencias, posibilitando a los usuarios encontrar información.

4.2.3. Tesauros

Un tesauro es un conjunto de términos formalmente controlado y estructurado que establece entre ellos relaciones semánticas genéricas. Los términos de un tesauro corresponden a una colección de ítems, de tal forma que un único tema representado mediante conceptos se describe por medio de descriptores. El tesauro incorpora todos los términos en una base de datos y cada uno de ellos se convierte en un punto de acceso para recuperación de la información (Blanco, 2007).

La potencia de un tesauro radica en la posibilidad de combinar todos los términos, convirtiéndolo en un lenguaje rico y potente, que resultará

muy útil para el análisis, descripción y recuperación automática de la información. La estructura de un tesauro está formada por:

- Unidades léxicas: existen dos tipos de unidades léxicas.
 - a. Descriptores: formado por los términos permitidos en el tesauro. Son un conjunto de palabras del lenguaje natural que expresan un concepto. Cada concepto se expresa mediante un único término y ese término corresponde a un único concepto. Distingue entre términos simples y compuestos.
 - b. No descriptores: para eliminar la sinonimia del lenguaje natural, se elige un único término de entre todos aquellos que son sinónimos y se rechaza el resto. Los no descriptores son precisamente el conjunto de términos desecharados.
- Relaciones semánticas: son aquellas relaciones que se establecen entre los términos de un tesauro. Los tesauros incorporan tres tipos de relaciones semánticas:
 - a. Equivalencia: proximidad fonética entre los descriptores y los no descriptores.
 - b. Jerárquicas: relaciones asimétricas establecidas entre los descriptores.
 - c. Asociativas: asocia descriptores que comparten lexema común.

4.2.4. Ontologías

El término ontología proviene de la filosofía y puede definirse como la rama correspondiente a la metafísica que se encarga del estudio de la naturaleza de la existencia. Este término se ha adoptado en la computación y Gruber lo define como una especificación explícita y formal de una conceptualización compartida (Gruber, 1993).

Para explicar esta definición describiremos por un lado, el sentido de la conceptualización, el cual se refiere al modelo abstracto del dominio de aplicación específico que se representa en la ontología, mientras que el término explícita se refiere a la necesidad de especificar los diferentes conceptos. Por otro lado, el aspecto formal se refiere a que debe indicarse la especificación mediante un lenguaje formalizado y la característica compartida alude al hecho de que una ontología debe representar el conocimiento aceptado por los usuarios que van a utilizarlo.

Otra definición de ontología fue dada por Weigand (Weigand, 1997; Staab y Studer, 2004), que la caracteriza como una base de datos donde se describen los conceptos de un dominio, sus propiedades y las relaciones entre ellos.

Podemos afirmar pues que una ontología no sólo formaliza un dominio, sino que el conocimiento que se representa en ella puede ser reutilizado. Las clases identifican los conceptos representativos del dominio, mientras que las propiedades se refieren a las relaciones entre los diferentes conceptos. Conocidas las clases y las propiedades es

necesario representar en la base de conocimiento instancias de las mismas.

Las ontologías también permiten definir reglas para modelar el conocimiento que puede no ser representado por las clases, propiedades e instancias. Estas reglas tienen que ser fácilmente interpretadas por las personas y deben ser utilizadas de forma eficiente en cálculos computacionales. Las reglas pueden definirse para:

- a) Creación de conocimientos: permite obtener nuevas sentencias lógicas, a partir de la información almacenada en la base del conocimiento.
- b) Restricciones: indican las propiedades que debe cumplir el modelo formalizado, detectando inconsistencias del mismo.
- c) Reglas reactivas: determinan acciones que debe tomar el sistema como consecuencia de cumplir ciertas condiciones.

Una ontología por tanto, no es un compendio del conocimiento objetivo, sino que depende del punto de vista de un grupo, por ello la ontología se compone de:

- a) Conceptos que describen el dominio.
- b) Definiciones de los conceptos.
- c) Relaciones entre los conceptos. Las relaciones pueden ser de varios tipos: taxonomía (es un), meronimia (es parte de), asociaciones, etc.
- d) Instancias que se utilizan para representar objetos determinados de un concepto.

4.3. Ontologías turísticas

El turismo es un ámbito de conocimiento que ha crecido espectacularmente en Internet, produciendo un gran cambio en la forma en que los turistas acceden a la información sobre los destinos, formas de contratar o reservas viajes, hoteles, etc.

La cantidad de información existente es ingente, que además de no estar normalizada y centralizada, se hace por tanto muy compleja de gestionar. Uno de los principales obstáculos que se encuentran los turistas es la dificultad de localizar la información completa y actualizada que demanda en un instante determinado.

Para poder gestionar toda la información disponible existente sobre los recursos (que constantemente va creciendo), surge el concepto de web semántica, que pretende proporcionar un significado bien definido a la información, teniendo como última finalidad facilitar la comunicación y el intercambio de información entre diferentes sistemas y entidades.

No sólo hay que compartir, intercambiar e integrar la información de los diferentes agentes de la industria turística, sino que se deben ofrecer también servicios turísticos personalizados (Barta et al., 2009; Herrero et al., 2012).

En el ámbito turístico, se han realizado varios estudios sobre taxonomías y ontologías relacionadas con el dominio del turismo, por parte de centros de investigación, universidades y la propia industria turística, con objeto de facilitar la gestión y el intercambio de la

información. En las siguientes secciones vamos a presentar algunas de las ontologías turísticas existentes.

4.3.1. Harmonise

Se trata de un proyecto de la Unión Europea para mejorar la interoperabilidad, en el que participaron varios estados y entidades. Su principal objetivo es el intercambio de datos entre diferentes entes, creando una red internacional (HarmoNET) que permitiera la normalización e intercambio de información en la industria turística (Barta et al., 2009).

Este proyecto pretende proporcionar a las organizaciones turísticas la capacidad de intercambiar información, sin tener que modificar sus estructuras internas ni sus sistemas de información. Para ello, hubo que establecer un conjunto mínimo de conceptos que representasen las ideas más relevantes del turismo, desarrollando la ontología IMHO que identifica y define los conceptos turísticos más relevantes (Siricharoen, 2010).

Harmonise actúa como una ontología mediadora entre diferentes ontologías turísticas, realizando una función de enlace semántico entre distintos sistemas, que permiten al receptor interpretar las fuentes de información como una extensión de sus propios datos sin preocuparse de cómo están organizados. La ontología está implementada mediante el lenguaje RDFS⁵⁶, conteniendo aproximadamente 200 conceptos basados en el campo del alojamiento, eventos, gastronomía, monumentos y lugares de interés.

⁵⁶ Resource Description Framework Schema

4.3.2. Mondeca

Es una ontología privada, por lo que no se permite su libre utilización. incluye conceptos definidos en el tesoro de la OMT (Organización mundial del turismo). Se desarrolló utilizando el lenguaje OWL⁵⁷ y contiene aproximadamente 1000 conceptos. Proporciona soluciones para la creación de taxonomías y gestión de ontologías (Siricharoen, 2010).

4.3.3. Hi-Touch

Se trata de una ontología desarrollada dentro del programa europeo IST/CRAFT Hi-Touch, destinado al desarrollo de metodologías basadas en web semánticas, con el objetivo de formalizar el conocimiento sobre las expectativas del turista, para de esa forma proporcionar servicios turísticos personalizados (Euzénat, Remize y Ochanine, 2003).

Está diseñada utilizando el lenguaje OWL y se basa en clasificar las entidades turísticas relacionándolas a partir de una red semántica. Los ítems turísticos son clasificados por palabras clave utilizando el tesoro de la OMT, garantizando la consistencia de la categorización de los recursos turísticos.

4.3.4. QALL-ME

Forma parte de un proyecto financiado por la Unión Europea, cuyo objetivo es proporcionar una infraestructura que permita que los usuarios realicen cuestiones en lenguaje natural en diferentes idiomas, para que el sistema devuelva una lista de respuestas, teniendo en cuenta la modalidad seleccionada por el usuario (Ou et al., 2008). Abarca

⁵⁷ Web Ontology Language

conceptos relacionados con destinos, alojamientos, puntos de interés, eventos y medios de transporte. Está escrita utilizando el lenguaje OWLDL, compuesta por más de 120 clases y más de 100 relaciones existentes entre ellas.

4.3.5. DERI

Incluida en el proyecto ONTOUR de la universidad de Innsbruck, tiene como objetivo crear un buscador web turístico basándose en la web semántica (Siorpaes y Bachlechner, 2006). Está escrito en el lenguaje OWL y se centra en la descripción de alojamientos e infraestructura que permite al usuario encontrar la información deseada. Se basa en el tesauro de la OMT y una de sus principales aportaciones es la utilización de aspectos geográficos como las coordenadas de un punto turístico, para facilitar el cálculo de distancias entre distintos puntos.

4.3.6. EON

Se trata de una ontología desarrollada en Francia que permite describir los conceptos turísticos clasificados en categorías (Barta et al., 2009). Incluye alojamientos, medios de transporte y monumentos. Es una de las primeras ontologías turísticas, aunque actualmente está en desuso.

4.3.7. cDott

Desarrollada por la universidad de Viena, se basa en la ontología Harmonise, teniendo como objetivo facilitar la interoperabilidad en operaciones a bajo nivel, que permita proporcionar una estructura de varias ontologías específicas interconectadas entre sí, para así extender

su alcance y funcionalidad, estando desarrollada en OWL DL (Barta et al., 2009).

4.3.8. Otras ontologías

En la presente sección se hace una enumeración de otras ontologías menos conocidas:

- a) TAGA: se centra en el ámbito de itinerarios de viajes, servicios de reservas, etc. Está escrita en el lenguaje RDF y OWL (Zou et al, 2003).
- b) CRUZAR: desarrollada en un proyecto del ayuntamiento de Zaragoza, se trata de una ontología desarrollada para la construcción de una aplicación, para la obtención de rutas turísticas a partir del perfil y contexto del turista. Se basa en capturar información de tres entidades: recursos turísticos, perfiles del turista y rutas turísticas (Minguez, Berrueta y Polo, 2009).
- c) GETESS: ontología basada en desarrollar una aplicación que permite recuperar información relevante de páginas web turísticas, para una posterior clasificación e interpretación semántica. Consta de más de 1000 conceptos y 200 relaciones (Staab y Studer, 2004)
- d) ebSemantics: creada por la fundación AUSTRIAPRO, está desarrollada en el lenguaje OWL y la información que gestiona está relacionada con los alojamientos, eventos y gastronomía (Barta et al., 2009).

4.4. Escenarios para el desarrollo de ontologías

El conjunto de aplicaciones y técnicas que queremos desarrollar para los turistas requieren la existencia de ontologías. Existen varias metodologías para su desarrollo, siendo sus principales limitaciones la imposibilidad de desarrollarse de forma cooperativa y la no reutilización de recursos ontológicos ya existentes (Suárez-Figueroa y Gómez-Pérez, 2008). Existen varias formas diferentes de crear las ontologías:

- a) Escenario 1: comenzar desde cero.
- b) Escenario 2: construcción reutilizando recursos no ontológicos.
- c) Escenario 3: construcción reutilizando recursos ontológicos.
- d) Escenario 4: construcción reutilizando y aplicando procesos de reingeniería de recursos ontológicos.
- e) Escenario 5: construcción reutilizando e integrando recursos ontológicos.

La mayoría de las ontologías comentadas en las secciones anteriores han seguido el escenario primero, es decir, han sido desarrolladas desde cero, aunque es factible el desarrollo de una red de ontologías mediante una combinación de escenarios.

4.4.1. Especificación de la ontología

El subprocesso de especificación de las ontologías consiste en la recogida de los diferentes requerimientos que la ontología debe cumplir. Para este propósito deben recogerse todos los requerimientos en un documento denominado ORSD (Documento de especificación de

requerimientos de la ontología), donde deben incluirse de forma detallada:

- a) Propósito de la ontología.
- b) Nivel de formalidad.
- c) Alcance.
- d) Usuarios previstos.
- e) Tipos de usos que se le va a dar.

Noy y MacGuinness (Noy y MacGuinness, 2001) proponen los siguientes pasos para recoger los requerimientos de la ontología:

- a) Intención para la cual se va a utilizar.
- b) Usos que se le dará.
- c) Exigencias que tiene que satisfacer después de su implementación.

Existen diferentes técnicas para realizar la tarea indicada anteriormente, entre las más comunes destacan:

- a) Sesiones de brainstorming o tormentas de ideas entre los desarrolladores, usuarios, ingenieros, etc.
- b) Puesta en escena de los casos de uso en los que se utilizará la ontología.
- c) Preguntas de competencia (Uschold y Grüninger, 1995), mediante cuestiones basadas en lenguaje natural que la ontología debe responder una vez que se ha desarrollado.

La metodología NeOn propone las siguientes recomendaciones para obtener los requerimientos de las redes de ontologías (Suárez-Figueroa y Gómez-Pérez, 2009):

- a) Definir el objeto, alcance y nivel de formalización.
- b) Identificar los usuarios y usos potenciales.
- c) Clarificar los requerimientos, agrupándolos, validándolos y priorizándolos.
- d) Extraer la terminología y la frecuencia.

4.4.2. Ciclo de vida de una ontología

Una vez obtenidas las especificaciones de los requerimientos, se deben identificar y programar el resto de actividades que se tienen que llevar a cabo en el proceso de desarrollo de la ontología:

- a) Determinar el modelo de ciclo de vida de la red de ontologías.
- b) Seleccionar los procesos y las actividades.
- c) Establecer las relaciones entre los procesos y las actividades.
- d) Asignar la secuencia y prioridad de las actividades y procesos.
- e) Identificar restricciones de recursos y tiempo.

Las ontologías se pueden desarrollar en diferentes lenguajes como por ejemplo RDF, XML SCHEMA y OWL. Este último es el más utilizado por tener una mayor capacidad para expresar significado y semántica. Los editores ontológicos permiten comprender el dominio del concepto de forma sencilla y clara. Entre ellos destaca el editor PROTEGE, creado por la Universidad de Standford, que está desarrollado sobre plataforma libre. PROTEGE permite visualizar, editar clases y propiedades de la ontología de forma cooperativa. Además soporta la

edición de comentarios y notas, permitiendo la creación de debates a lo largo del ciclo de vida del desarrollo ontológico.

4.5. Organización del perfil del usuario

Hasta ahora hemos descrito el estado del arte sobre cómo poder representar el conocimiento de los ítems, pero no hemos profundizado en la manera en la que se modelaría la información del usuario que utilice el sistema que propondremos en capítulos posteriores. Los perfiles de usuario son un elemento indispensable en los sistemas de recomendación, ya que nos permitirá conocer los gustos y preferencias del mismo. El tener toda esta información correctamente registrada redundará en el resultado de las recomendaciones generadas.

Existen varios mecanismos o métodos para generar los perfiles de usuario y para mantenerlos, entre los que destacamos:

- a) Representar el perfil del usuario.
- b) Crear perfiles iniciales.
- c) Retroalimentar los perfiles del usuario.
- d) Adaptar el perfil del usuario ante cambios significativos de las preferencias del mismo.

4.5.1. Representar información del perfil del usuario

Disponer del perfil del usuario es indispensable en el proceso de personalización de un sistema de recomendación. Si el perfil está adecuadamente construido, el sistema irá realizando recomendaciones que satisfagan al usuario, consiguiendo de esta manera aumentar su confianza en el sistema. Existen varios tipos de técnicas para

representar la información sobre el perfil de los usuarios, entre ellas podemos destacar:

- a) Historial: consiste en almacenar registros de los ítems que el usuario ha utilizado o visitado en el pasado, así como el nivel de valoración que estos ítems producen en el usuario. Existen muchos ejemplos de sistemas de recomendación basados en registro de historial, como puede ser Amazon. Algunos sistemas utilizan varias listas de registros: una con los ítems que realmente ha utilizado y otra sobre los ítems que no le interesan (Blanco et al., 2006).
- b) Vectores de características: cada ítem utilizado por el usuario se representa mediante un vector de características. Cada elemento del vector almacena un valor que representa la relevancia de una determinada característica relacionada con el ítem. De esta forma, dos ítems son similares si sus vectores de características son muy parecidos. Un ejemplo de este tipo de modelado lo encontramos en MovieLens (Herlocker et al., 2004).
- c) Matrices: algunos sistemas utilizan para almacenar los perfiles de usuario las matrices. No es un buen método para aquellos sistemas de recomendación que tengan muchos usuarios o ítems a recomendar, ya que el número de filas de la matriz coincide con el número de usuarios, mientras que el número de columnas será igual al número de ítems a recomendar. Así, el elemento de la fila i , columna j almacena el nivel de interés que le ha producido al usuario i , el ítem j . En caso que el usuario no haya utilizado el

ítem j, el valor almacenado será un valor nulo. Un ejemplo lo encontramos en Smart Radio (Colucci et al., 2003).

- d) Datos demográficos: se almacena en el perfil del usuario información personal como sexo, edad, estado civil, profesión, etc. Estas características demográficas permiten identificar al tipo de usuario y de esta forma recomendar ítems que han sido valorados positivamente por usuarios del mismo tipo. La información que se proporciona al sistema es indicada por el propio interesado de forma explícita. El principal inconveniente es que los usuarios no siempre están de acuerdo en revelar información personal, así como tener que llenar formularios bastante completos y minuciosos antes de comenzar a utilizar el sistema. Un ejemplo que utiliza este tipo de perfil es Blu-IS (Bartfai, 1994), que complementa la información demográfica de los usuarios con atributos contextuales como lugar, hora e historial anterior.
- e) Redes semánticas: se utilizan técnicas para constituir los perfiles de los usuarios basadas en representar el conocimiento mediante redes semánticas. Se representan los conceptos que identifican las preferencias en nodos de una red y las relaciones entre ellos mediante enlaces, además de asociar valores numéricos que identifican el nivel de interés del usuario en un concepto determinado.
- f) Redes bayesianas: las redes bayesianas son unos grafos acíclicos donde los nodos representan variables proposicionales y los arcos las dependencias entre nodos (Jensen, 2001). El valor de cada nodo depende de los valores de los nodos padres calculados a

partir del teorema de Bayes. El perfil del usuario puede utilizar una red bayesiana para representar las preferencias del usuario, de forma que los nodos y los arcos identifican los conceptos que interesan a los usuarios y las relaciones causales entre ellos.

- g) Árboles de decisión: se trata de una estructura jerárquica, donde los nodos identifican preguntas sobre las variables consideradas en el problema de clasificación, mientras que los arcos representan las respuestas y posibles valores de las variables anteriormente indicadas.
- h) Aprendizaje inductivo: utiliza reglas de asociación de forma que permiten identificar patrones de recomendación.
- i) Redes neuronales: son sistemas de aprendizaje y procesamiento automático de la información (Arbib, 1995), que se encargan de crear una red neuronal por usuario, para poder aprender y representar sus preferencias.
- j) Ontologías: la forma más exhaustiva de formalizar el conocimiento (tal y como hemos visto en secciones anteriores) es la ontología, de forma que se identifican todos los conceptos y relaciones del dominio. Mediante la ontología es posible mantener de forma eficiente un modelo semántico completo sobre las preferencias de los usuarios.

4.5.2. Creación del perfil inicial del usuario

Existen varias técnicas para la construcción del perfil inicial del usuario. Las principales son:

- a) Reconocimiento automático: algunos sistemas no requieren de un perfil inicial, sino que sus preferencias se van creando a través de la interacción del usuario con el sistema. Generalmente se utiliza en sistemas de personalización web.
- b) Inicialización explícita: los usuarios deben describir sus intereses y preferencias para construir su perfil inicial. La principal ventaja es que se dispone de una precisión adecuada, pero el inconveniente es que el usuario tiene que asumir que para poder utilizar el sistema, debe llenar varios formularios que pueden hacerse tediosos.
- c) Estereotipos: consiste en definir categorías a las que pertenece el usuario en función de preferencias y datos demográficos. De esta forma, cuando un usuario se registra proporciona información sobre él y automáticamente se le asigna un estereotipo inicial.
- d) Entrenamiento: se basa en presentar al usuario un conjunto de ítems, para que indique claramente lo que le interesa y lo que no. El principal inconveniente es que los ejemplos que se toman deben ser suficientemente representativos y conocidos por todos los usuarios.

4.5.3. Retroalimentación

Una vez que el usuario va interactuando con el sistema, la información que proviene de las evaluaciones de las recomendaciones realizadas es una fuente de conocimiento muy útil, gracias a esta información, el sistema puede modificar los intereses del usuario para converger a los intereses reales del mismo. Existen varias técnicas:

- a) Sin retroalimentación automática: este tipo de técnica delega en el propio usuario para que actualice su perfil, cuando considere que los resultados no se adecuan a sus intereses reales.
- b) Retroalimentación explícita: una vez se le recomienda algún ítem, debe declarar el nivel de interés que le ha producido el ítem recomendado.
- c) Retroalimentación implícita: se lleva a cabo a partir de las acciones e interacciones que realiza el usuario con el sistema.
- d) Retroalimentación híbrida: combinación de las técnicas anteriores.

4.5.4. Adaptación del perfil

Las preferencias de un usuario pueden cambiar durante el tiempo, por lo que una adecuada adaptación del perfil del usuario debe llevarse a cabo regularmente, para permitir la aparición de nuevas preferencias, etc. Es evidente que la retroalimentación es un elemento clave, pero las técnicas que pueden utilizarse para adaptar el perfil son:

- a) Actualización manual del perfil: el usuario debe modificar su propio perfil cuando sus preferencias hayan cambiado.
- b) Incorporación de nueva información al perfil: la retroalimentación va modificando directamente el perfil, pero no tiene capacidad de olvido.
- c) Olvido gradual: consiste en considerar o darle más peso a los datos más reciente, haciendo que los más antiguos tengan un valor más relativo. El inconveniente está en decidir a partir de qué momento se debe realizar un olvido total.

Capítulo 5

RAMCAT. Un marco de recomendación en un sistema de realidad aumentada aplicada al Turismo

5.1. Introducción

En este capítulo vamos a presentar las bases del marco propuesto y su aplicación en un sistema de realidad aumentada de un destino turístico. Describiremos sus componentes, sistemas y arquitectura. La propuesta está basada en una ontología que nos permitirá representar información sobre los turistas, contexto y puntos de interés turísticos.

El marco propuesto necesita de unas definiciones y formalismos básicos, relacionados con los modelos de recomendación utilizados y con

la metodología general que permitirá obtener los resultados. Los objetivos de la tesis se centran en la inclusión de atributos contextuales en los sistemas de recomendación y su aplicación en un sistema de realidad aumentada de un destino turístico.

La principal ventaja que puede aportar la conjunción de técnicas de recomendación y de realidad aumentada es que pueden complementarse para facilitar la experiencia del turista que visita un destino. El uso de técnicas de recomendación permitirá que el sistema de realidad aumentada muestre solo aquella información que es realmente relevante para el turista, mientras que la aplicación de técnicas de realidad aumentada permitirá no sólo una mejor difusión y promoción del destino, sino que ayudará eficientemente al turista en la búsqueda de puntos de interés, así como a guiarlo adecuadamente en el destino.

Se trata por tanto, de una clara apuesta por la interoperabilidad e integración real de diferentes sistemas, que facilitará la interacción entre los turistas y el destino. La innovación no es solo adoptar nuevas tecnologías, sino integrarlas.

Se pretende por tanto crear un modelo que mejore las técnicas existentes de recomendación y aplicarlo a un sistema de realidad aumentada. El proceso de recomendación permitirá mostrar resultados tanto para individuos como para grupos, utilizando atributos contextuales y teniendo en cuenta las necesidades y características del usuario que realiza la petición de recomendación. En determinadas situaciones, puede resultar muy interesante que el turista indique algunas características de lo que quiere buscar, así como que especifique

algún ítem ya visitado con anterioridad para que se le recomienden ítems similares.

Nos podemos encontrar con una casuística muy diversa: turistas que viajan solos, que viajan en grupo, que no han utilizado el sistema nunca, que han visitado y valorado muchos puntos turísticos en el sistema, etc. El sistema debe tener en cuenta estas circunstancias y responder adecuadamente a las distintas demandas del turista que lo utiliza.

El capítulo comienza con una descripción de las principales ventajas que pueden aportar los sistemas de realidad aumentada a un destino turístico, enumerando los distintos escenarios de uso del marco propuesto.

Posteriormente, se detallarán los diferentes sistemas de la arquitectura RAMCAT. Además, se aporta la descripción de la ontología utilizada para representar el conocimiento de los puntos turísticos y de los turistas.

Seguidamente, se definirán diferentes procesos y modos de trabajo del sistema, así como interacciones y tipos de filtrado para poder obtener las recomendaciones. Para finalizar, se expondrán las actividades realizadas por el turista en su utilización del sistema, los mecanismos de evaluación de ítems visitados y el funcionamiento general del sistema, mediante el uso de los diferentes motores de recomendación que se describen en el capítulo 6.

5.2. Aplicación de la realidad aumentada a un destino turístico

Cada día es más importante el uso de las nuevas tecnologías en actividades relacionadas con el turismo. Dentro de éstas, la realidad aumentada aparece como una alternativa relevante para la mejora de los servicios que ofertan los destinos turísticos. Aunque en el capítulo 2 se ha realizado un estudio del estado del arte, en esta sección queremos profundizar en aquellos aspectos relacionados con el turismo donde nuestra propuesta de aplicación de modelos contextuales de recomendación en un sistema de realidad aumentada tiene especial interés. En concreto, podemos indicar que este tipo de tecnología representa un método eficiente e innovador para el turismo en actividades de:

- a) Difusión
- b) Promoción
- c) Guiado
- d) Búsqueda de puntos turísticos
- e) Organización del viaje

A continuación presentamos una descripción de cada una de estas áreas de aplicación de la realidad aumentada al turismo, así como sus potenciales aportaciones a los destinos turísticos.

5.2.1. Difusión del patrimonio de un destino turístico

La difusión del patrimonio es una actividad que los destinos turísticos se encargan de realizar para dar a conocer sus monumentos y su entorno. Se trata sin duda de una actividad que permite fomentar y

aumentar el número de visitas a estos puntos de interés. La realidad virtual y la realidad aumentada permiten realizar aportaciones muy interesantes en la difusión del patrimonio (Xiangn y Pan, 2011; Liburd, 2012), entre las que podemos destacar:

- a) Reconstrucción de monumentos desaparecidos total o parcialmente. La realidad aumentada puede mostrar al turista información sobre edificios o monumentos que están total o parcialmente destruidos, permitiendo presentar una recreación virtual o una combinación de los restos existentes con una reconstrucción parcial de los elementos desaparecidos (Choundary et al., 2009).
- b) Visualización de elementos de difícil acceso. Muchas veces se visitan monumentos donde existen detalles arquitectónicos que el turista no puede ver porque resultan inapreciables desde las localizaciones de acceso de los mismos. También, nos podemos encontrar dificultades de movilidad del turista, que le impida acceder a algunos elementos. En estos casos la realidad aumentada permitirá mostrar a los turistas esos elementos de información (Henderson et al., 2009).
- c) Recreación de entornos de acceso restringido. Determinados puntos de interés tienen un acceso restringido, ya que un número excesivo de visitas puede hacer que su conservación se ponga en peligro. Para solventar esta restricción se utilizan habitualmente reconstrucciones o maquetas físicas. La recreación virtual resulta más económica y ofrece muchas alternativas de interacción (Hsu, 2011; Leiva, Guevara y Rossi, 2012).

- d) Evolución histórica de un monumento. Este tipo de tecnologías posibilita que el turista pueda conocer la evolución en el tiempo de un monumento, permitiendo obtener una representación del mismo en diferentes momentos históricos (Chen et al., 2011, Leiva et al., 2012).
- e) Interpretación de un monumento. La visualización de un monumento puede complementarse con información virtual sobre su uso y escenas pertenecientes a momentos históricos, ofreciendo al turista una información dinámica y realista (Xiang y Pan, 201; Liburd, 2012). Un ejemplo de este tipo de aplicación consiste en complementar la imagen de edificio histórico con mobiliario de la época, recreación de escenas con personajes históricos, etc.
- f) Interacción virtual. Resulta muy interesante que al turista se le ofrezca información interactiva sobre el punto de interés que está visitando (Hsu, 2011), siendo muy habitual la utilización de asistentes que guían al turista durante la visita.
- g) Creación de rutas. Cuando un turista visita una zona en la que existen varios puntos patrimoniales interesantes, la posibilidad de disponer de una ruta de visita adaptada a las preferencias del turista, así como una ayuda para guiarle, resultará especialmente atractivo (Leiva, Guevara y Rossi, 2012).
- h) Modernización de información de promoción turística. En muchas oficinas de turismo suelen aparecer paneles con fotografías de monumentos y otros tipos de recursos muy diversos, que permiten a los turistas obtener información de la oferta del destino (Troutiño et al., 2011). Las posibilidades de la realidad

aumentada para atraer a los visitantes y ofrecerles una plataforma más actual para completar las necesidades de información son muy interesantes (Leiva, Guevara y Rossi, 2012).

- i) Información sobre intervenciones. Muchas veces los turistas no pueden visitar completamente un monumento porque está siendo restaurado, siendo esto un elemento de decepción de la visita. Las herramientas de realidad aumentada permiten ofrecer al turista información de la intervención que se está haciendo e incluso una recreación virtual del resultado final, lo que representa un importante valor añadido (Leiva, Guevara y Rossi, 2012).

5.2.2. Promoción del destino

La promoción y la difusión del destino son dos conceptos muy relacionados. La difusión está asociada a la transmisión y presentación a los turistas de los distintos puntos y actividades turísticas del destino, mientras que la promoción tiene un matiz publicitario y de venta. La promoción de un punto de interés turístico (monumento, parque, restaurante, teatro, etc.) tiene como objetivo atraer a los turistas. El tipo de aplicaciones de las tecnologías de realidad aumentada a la promoción de un destino son similares a las descritas en el apartado anterior:

- a) Folletos de promoción. Los folletos existentes en las oficinas de turismo del destino deben aportar los elementos más novedosos a los turistas (Fesenmaier, Werthner y Wober, 2006). La utilización de técnicas de realidad aumentada puede ser una

forma estimulante de dirigirse a los visitantes en las oficinas de turismo, hoteles, monumentos, etc.

- b) Promoción mediante la recreación virtual de actividades de interés cultural. El uso de la realidad aumentada para representar actividades típicas de un entorno, permite promocionar y fomentar el uso de estas actividades (Leiva, Guevara y Rossi, 2012).

5.2.3. Guiado

La utilización de la tecnología basada en realidad aumentada es considerada un elemento fundamental y eficiente para guiar al turista (Henderson et al., 2009). El guiado tiene como objetivo básico ayudar al visitante a localizar puntos de interés, consiguiendo que el turista sienta que su visita está siendo provechosa en todo momento, no perdiendo más tiempo del necesario en desplazamientos y en visitas menos importantes. Podemos encontrar diferentes aplicaciones orientadas al guiado (Leiva, Guevara y Rossi, 2012):

- a) Mapas. La realidad aumentada permite dar información rápida y bastante precisa sobre los lugares donde se encuentran determinados puntos turísticos, distancias desde la posición en la que está el turista, distancia entre diferentes puntos, etc.
- b) Información sobre medios de transporte. Este tipo de tecnología representa una interesante herramienta para poder informar a los turistas sobre los lugares donde se encuentran los medios de transportes necesarios para visitar un determinado ítem, así como los tiempos de espera asociados. De esta forma, los turistas

entienden el destino como una conjunción no sólo de recursos y actividades, sino de mecanismos para llegar a ellos optimizando su visita.

- c) Traductores. Uno de los principales problemas del guiado es que la información que se encuentra en paneles, señales, etc., con frecuencia está escrita en un idioma desconocido para el turista. Existen aplicaciones de realidad aumentada que son capaces de traducir y por tanto conseguir que el turista esté en todo momento informado del lugar en el que se encuentra.
- d) Información sobre intervenciones de tráfico. A veces la existencia de obras en un destino puede provocar a los turistas cambios en la visita que tenía planificada inicialmente. La realidad aumentada aparece como una alternativa que podría informar al turista de alternativas (nuevos accesos peatonales, nuevas paradas de autobuses, etc.) que minimicen las molestias que pueda sufrir.

5.2.4. Búsqueda de puntos turísticos

Es muy habitual que los turistas visiten oficinas de turismo del destino para obtener información sobre puntos de interés destacados en la zona. Disponer de un asistente que le permita buscar qué visitar, informándole de horarios y precios, constituye una importante ayuda en la búsqueda de información (Fesenmaier, Werthner y Wober, 2006). Poseer herramientas de realidad aumentada basadas en sistemas de recomendación es una técnica que los destinos pueden adoptar para facilitar la búsqueda de los puntos que el turista debe visitar en el

destino. Entre las diferentes opciones de búsqueda que un sistema de realidad aumentada debería permitir destacamos:

- a) Búsqueda por categorías. El turista selecciona el tipo de recursos que está interesado en visitar. De esta manera informa de sus intenciones, mostrando el sistema los puntos que cumplen sus expectativas, distancias, etc.
- b) Recomendación. Las herramientas basadas en sistemas de recomendación son capaces de recomendar aquellos puntos de interés que cumplirán con las preferencias del turista.

5.2.5. Organización del viaje

Las tecnologías basadas en realidad aumentada pueden mostrar información de muy diverso tipo de los diferentes puntos o recursos turísticos. La posibilidad de que estas herramientas puedan además organizar el viaje es una importante ayuda. Estos sistemas deben basarse en técnicas de recomendación, que mediante diferentes algoritmos, permitan establecer una ruta con los diversos puntos y actividades teniendo en cuenta las preferencias del turista, así como el tiempo estimado para la realización de la visita, costes, etc. De esta forma, la realidad aumentada nos permitirá mostrar información de la ruta basándose en herramientas de guiado.

5.2.6. Funcionalidad de la propuesta.

La realidad aumentada, como se ha definido en el capítulo segundo, es una tecnología que superpone información virtual sobre la información real (Azuma et al., 2001). El uso de la realidad aumentada ofrecerá diferentes funcionalidades en situaciones de movilidad del

turista. Esto no significa que el visitante no pueda planificar o informarse con el sistema de realidad aumentada antes de realizar la visita, pero es evidente que la aplicación de esta técnica durante la misma enriquecerá su experiencia.

Las principales características del uso de la realidad aumentada en movilidad son:

- a) Proporciona una escena que combina información real con virtual, visualizándola en tiempo real.
- b) Permite conocer la posición y orientación desde el punto de vista del turista en todo momento.
- c) Los objetos virtuales se muestran en posición, orientación y a escala en el entorno real, pudiendo observarse desde diferentes posiciones.

El sistema de realidad aumentada propuesto (RAMCAT) permite mostrar información al turista sobre actividades y puntos de interés turísticos. Esta información se centra en el propio turista, teniendo en cuenta sus preferencias personales y diferentes atributos contextuales. El conjunto de información que puede mostrar incluye visitas a museos, monumentos, jardines, alojamientos, restaurantes, actividades culturales y de ocio, etc. El turista podrá obtener información detallada de las actividades recomendadas, así como solicitar rutas a partir de ellas.

Una de las principales aportaciones que tiene el sistema es que no sólo filtra la información basándose en preferencias y valoraciones de turistas con perfil similar, sino que tiene en cuenta también el uso que se hace del sistema y los atributos contextuales que puedan afectar a la

visita (clima, compañía, ...). RAMCAT irá recopilando información para mejorar las recomendaciones a partir de la interacción con el turista.

Para el registro del usuario y para obtener información sobre sus preferencias, resulta mucho más eficiente realizarlo mediante un sistema web que permitirá almacenar la información del turista, aunque también se debe permitir el registro desde la aplicación móvil. Una vez que el sistema envíe recomendaciones al turista, éste en su dispositivo móvil podrá elegir lo que quiere visitar, así como descargar la información necesaria para realizar la visita. La disponibilidad de información contextualizada a la posición del turista hace que el sistema deba poseer tecnologías de posicionamiento GPS.

La propuesta de utilizar la realidad aumentada junto a un sistema de recomendación contextual supone un importante elemento innovador en el ámbito turístico. El modelo que se propone permite que el turista pueda realizar diferentes actividades y actuar en diferentes escenarios como:

Actividad 1: Registro del turista en el sistema.

Actividad 2: Búsqueda por categorías.

Actividad 3: Solicitud de recomendación.

Actividad 4: Información de una actividad.

Actividad 5: Guiado de la actividad.

Actividad 6: Solicitud de ruta.

Actividad 7: Selección y configuración de motores de recomendación.

Actividad 8: Registro de visita en grupo.

Actividad 9: Acceso al perfil del turista.

Actividad 10: Editar atributos contextuales individuales.

Actividad 11: Evaluación de una actividad.

Este conjunto de actividades serán descritas y desarrolladas en el presente capítulo y en el siguiente.

5.3. Arquitectura RAMCAT

La arquitectura de un sistema define la estructura y el comportamiento del mismo. La arquitectura de esta propuesta se expresará en términos de entidades hardware, software y las relaciones entre ellas y, por tanto, quedará definida por:

- a) Estructura de los componentes.
- b) Relaciones entre los componentes.
- c) Relación con los usuarios y elementos externos.
- d) Principios y directrices que rigen el diseño.
- e) Evolución del sistema.

Una adecuada división de un sistema en componentes o subsistemas debe minimizar las comunicaciones entre éstos. Esto se ha tenido en cuenta al definir la arquitectura de RAMCAT, ya que requiere que los procesos computacionales se realicen tanto en el terminal móvil del turista como en el servidor del destino turístico. La figura 5.1 muestra los componentes hardware contemplados en la arquitectura, así como las interacciones entre ellos.



Figura 5.1. Arquitectura propuesta modelo RAMCAT

Como se puede observar, el turista accede al sistema a través del dispositivo móvil o de un ordenador, necesitando ambos conexión a Internet, mientras que el servidor realiza fundamentalmente una función de almacenamiento y de procesamiento de datos.

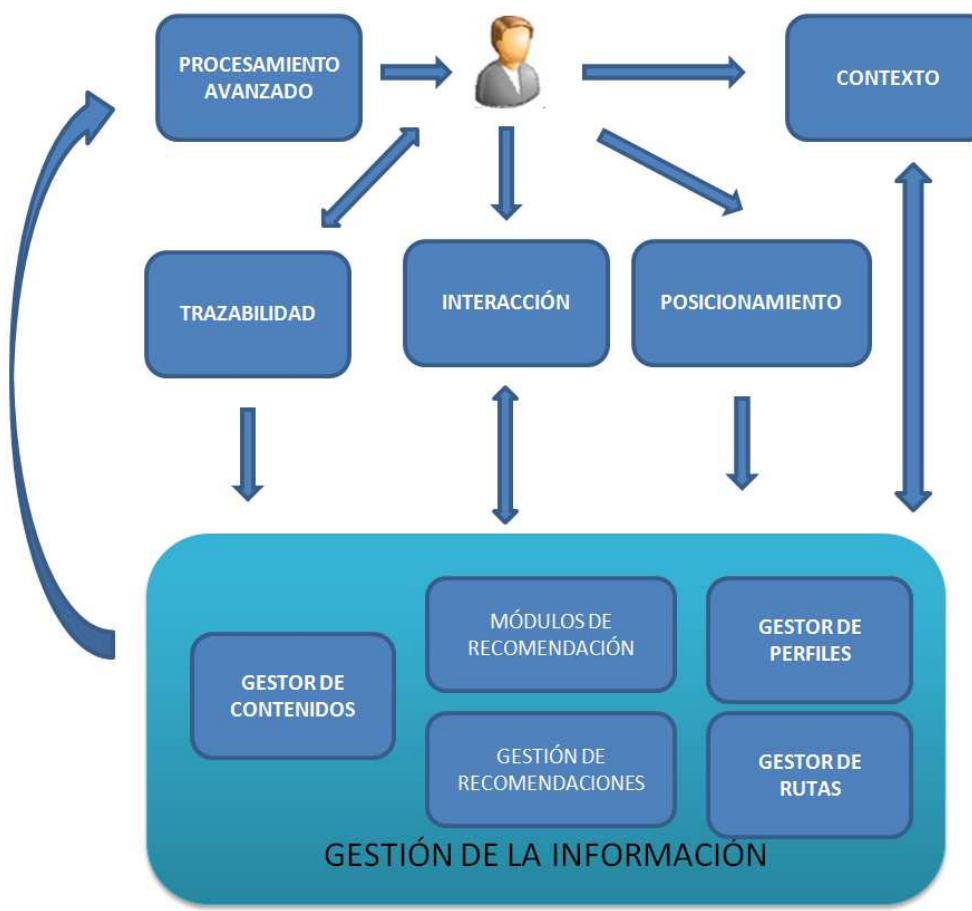


Figura 5.2. Componentes de la arquitectura RAMCAT

El comportamiento dinámico de esta arquitectura sería el siguiente: el punto de entrada del turista al sistema es a través de la aplicación móvil o de la aplicación web (formalmente lo denominaremos sistema de interacción o capa de interfaz), que posibilitan al turista, en un primer paso, registrarse en el sistema para posteriormente solicitar las recomendaciones. Este sistema recibe la información del turista y la convierte en diferentes eventos que son transferidos al sistema de gestión de la información (Figura 5.2). Otra fuente de entrada de información para el marco propuesto es el sistema de posicionamiento, que recoge información mediante GPS para determinar la posición del turista. El sistema de procesamiento avanzado es la parte de la

aplicación móvil encargada de presentar la información aumentada al turista. Este sistema recibe la información generada por el sistema de gestión de la información, en concreto del módulo de gestión de las recomendaciones, el cual las obtendrá partiendo de los diferentes motores de filtrado pertenecientes al módulo de recomendación. Una aportación destacada de la arquitectura propuesta es que para poder realizarse las recomendaciones, se utilizará información obtenida del sistema de contexto.

Para poder realizar las recomendaciones, los motores necesitan interactuar con el sistema gestor de contenidos, el sistema gestor de perfiles y el sistema de contexto. Cuando el turista solicita una ruta, ésta se generará mediante actividades obtenidas a partir del módulo de gestión de recomendaciones y del sistema de interacción. Éste último sistema permite al turista configurar la generación final de la ruta indicando diferentes parámetros como tiempos, costes, etc.

Respecto a la distribución física de los diferentes sistemas, se propone una *arquitectura mixta* en la que algunos sistemas se ejecutan en el servidor, otros en el dispositivo móvil y otros en ambos. Es evidente, que los motores de recomendación se deben ejecutar remotamente y suministrar los resultados a los dispositivos móviles, ya que la cantidad de cálculos computacionales requeridos es muy alta.

En las siguientes subsecciones se describen los componentes del sistema y el mecanismo para obtener información del turista para realizar las recomendaciones.

5.3.1. Componentes de RAMCAT

En la presente sección vamos a identificar y describir los diferentes módulos que componen el modelo propuesto. Los turistas que usen el sistema pueden utilizar dos tipos de interfaces: web y móvil (es decir, una aplicación nativa instalada en un dispositivo del tipo *smartphone*.) En este segundo caso, el dispositivo debe disponer de cámara, GPS y conexión inalámbrica. El sistema contará con un servidor del destino turístico, donde estará almacenada información sobre los puntos de interés del destino y de los usuarios que utilizan o han utilizado el sistema.

A partir de la arquitectura descrita, y para satisfacer las diferentes formas de interacción, se definen una serie de componentes o sistemas que implementarán eficientemente las funcionalidades requeridas en los diferentes escenarios de uso. Así, el sistema RAMCAT está compuesto por los siguientes componentes (véase la figura 5.2):

- Sistema de interacción o capa de interfaz: estará formado por dos subsistemas o aplicaciones, por un lado una aplicación web que fundamentalmente servirá para que los usuarios puedan registrarse, introducir datos sobre sus preferencias y consultar información sobre el destino; por otro lado, una aplicación móvil que permite el acceso del usuario al sistema mediante el dispositivo móvil permitiéndole consultar información, obtener recomendaciones, realizar valoraciones, etc.
- Sistema de gestión de la información: está compuesto, a su vez, por los sistemas de gestión de contenidos, de recomendación, de

gestión de recomendaciones, de gestión de perfiles y de rutas. En los siguientes apartados se describen estos sistemas detalladamente.

- Sistema gestor de contenidos de realidad aumentada: almacena toda la información sobre los puntos de interés, características, valoraciones, etc.
- Sistema de procesamiento avanzado: es el encargado de mostrar al turista la información de un ítem en forma de realidad aumentada, como, por ejemplo información de un cuadro o la recreación virtual de un punto de interés.
- Sistema de posicionamiento: obtiene la posición del turista a través de mecanismos de geolocalización (GPS).
- Gestor de perfiles de usuarios: almacena, controla y modifica toda la información del turista, distinguiendo entre diferentes perfiles para el mismo (preferencias individuales, de grupo, información recopilada a partir de visitas anteriores, etc.).
- Sistema de trazabilidad: tiene como objetivo conocer mejor al turista registrando los lugares visitados, peticiones realizadas, orden de visita de puntos turísticos, tiempo dedicado a cada actividad, aceptación o rechazo de recomendaciones, etc. Está directamente integrado con el sistema de interacción de usuarios, del que extrae información sobre la interacción del turista con el sistema, retroalimentando su perfil.
- Módulos de recomendación: cada uno de ellos es un sistema que aplicando diferentes técnicas de recomendación genera listas de puntos de interés adecuados al turista. Como característica

común, todos los módulos de recomendación utilizan atributos contextuales para obtener los valores de preferencia de los diferentes puntos de interés.

- Gestión de recomendación de puntos: recopila listas de recomendación generadas por los diferentes módulos de recomendación, obteniendo un listado ordenado por predicción de satisfacción del turista, utilizando técnicas de hibridación.
- Gestor de rutas: recoge del gestor de recomendación el listado de puntos con mayor valoración, teniendo en cuenta las preferencias y el contexto. A partir de estos puntos generará rutas basándose en el tiempo total de la visita, horarios, etc.
- Sistema de información de contexto: efectúa una recopilación de información contextual, realizando un pre-filtrado de ítems, basándose en este tipo de conocimiento. Como normalmente el número de puntos sobre los que habrá que calcular el posible interés de los turistas es bastante grande, este sistema atenúa este problema eliminando aquellos ítems que sean incompatibles con el contexto general y particular del turista que solicita la recomendación.

Como se puede comprobar, el sistema propuesto posee una estructura relativamente compleja para implementar las funciones de recomendación. No obstante, esta complejidad se simplifica en cierta medida ya que la arquitectura del sistema es:

- Flexible: como se describirá en el siguiente capítulo, actualmente está formado por cuatro módulos de recomendación, pero se le

pueden añadir nuevos motores de recomendación de forma independiente sin tener que modificar el resto del sistema.

- Personalizable: el turista puede habilitar y deshabilitar los diferentes módulos según la forma de interactuar o los resultados que deseé obtener.

A continuación se describen de forma más detallada cada uno de los sistemas.

5.3.2. Sistema de interacción o Capa de interfaz.

Este sistema está compuesto por dos aplicaciones: web y móvil. La aplicación web permite al turista registrarse, indicar sus preferencias, realizar búsquedas, solicitar recomendaciones, etc.

Respecto a la aplicación móvil, además de la funcionalidad de la aplicación web, permitirá mostrar información aumentada en un navegador de realidad aumentada. En concreto, se identifican dos tipos de interacciones del turista con el dispositivo móvil: interacción táctil con la pantalla del dispositivo e interacción basada en la posición del turista (aunque este tipo de interacción será implementada por el sistema de procesamiento avanzado, ya que el sistema debe posicionar automáticamente los elementos virtuales mientras el usuario cambia su ubicación u orientación). Este sistema es el encargado de proporcionar la interacción táctil usuario-máquina con el sistema, de forma que el turista puede seleccionar elementos interactivos para visualizar en su pantalla las diferentes funcionalidades ofrecidas por el sistema.

En este sistema el turista, una vez identificado, podrá indicar el modo de trabajo con el que va a utilizar la aplicación, cada uno de ellos asociado a una de las siguientes funcionalidades de RAMCAT:

- a) Modo recomendación: el sistema se basa en las preferencias del visitante para poder generar las recomendaciones.
- b) Modo búsqueda: el turista va a seleccionar directamente en el sistema actividades sin necesidad de recibir recomendaciones ni utilizar el perfil del turista.
- c) Modo aumentado: se visualiza en la pantalla la realidad aumentada apareciéndole por ejemplo un máximo de cinco puntos (esta cantidad será configurable) que son recomendados teniendo en cuenta el contexto, características y preferencias del turista que usa el sistema.
- d) Modo lista: se muestra una lista ordenada de puntos turísticos teniendo en cuenta las preferencias. El turista puede recorrer esta lista y obtener una descripción de los puntos recomendados. Una vez que selecciona el punto que desea visitar, podrá mediante el modo aumentado visualizar la imagen real con el punto seleccionado, lo cual le servirá para guiarse, informando, en el caso de que la distancia no sea cercana, de qué medios de transporte existen para llegar hasta ese punto.
- e) Modo ruta: en este modo de trabajo el turista selecciona de la lista aquellos puntos que desea visitar y el tiempo estimado que tiene para realizar la visita. El sistema le mostrará información sobre los puntos numerados con un orden recomendado de visita.

Esta ruta se estructura teniendo en cuenta las distancias, costes y horarios de las diferentes actividades.

5.3.3. Sistema de posición

Es muy importante disponer en todo momento de la posición y orientación del turista, para guiarlo correctamente al punto de interés seleccionado o para que le muestre las zonas donde se encuentran una serie de puntos recomendados. Mediante el dispositivo GPS se realizará un posicionamiento del turista que permitirá informar de las distancias a la que se encuentran los puntos turísticos.

El sistema debe proporcionar al sistema la información sobre la posición y orientación en tiempo real de la forma más precisa posible. Aunque son varias las tecnologías que podrían utilizarse para obtener y calcular esta información, incluso combinarse algunas de ellas, la tecnología más habitual y menos compleja es el uso de GPS en los dispositivos móviles utilizados por los turistas.

El sistema de posicionamiento global GPS usa señales enviadas por los satélites. Un cálculo correcto de la posición y vista real del turista posibilitará que la presentación de la información virtual esté contextualizada respecto a los objetos reales.

5.3.4. Sistema de procesamiento avanzado

Una vez que el turista en el sistema de interacción, decide la operación a realizar, el sistema debe responder adecuadamente a dichas peticiones mostrando, si es requerido, la información aumentada.

Por tanto, este sistema tiene como función principal descargar de trabajo al sistema de interacción encargándose exclusivamente de la parte relacionada con la realidad aumentada.

La información aumentada que se le presenta al turista consiste en la visualización en su dispositivo móvil de la imagen que la cámara captura en tiempo real, junto con los puntos de interés que se le ha recomendado visitar. En la figura 5.3 podemos ver una imagen del prototipo de interfaz desarrollado.



Figura 5.3. Pantalla del prototipo móvil RAMCAT

La información que el sistema podrá mostrar será de tipo multimedia, pues podrán integrarse diferentes tipos de datos texto, imágenes, audio,...). Esta información se le presenta al turista en la pantalla de su dispositivo móvil.

Otra función que debe permitir este sistema es visualizar información virtual en 3D. Esta contextualización de contenidos en la imagen real considerará la orientación del punto de vista del turista asociando el objeto a la escena real.

5.3.5. Sistema de gestión de la información

Se trata del elemento fundamental del marco propuesto. Es el responsable de recomendar a los turistas aquellos puntos de interés que estén en concordancia con las preferencias del mismo así como de su contexto. Este sistema se compone como se ha indicado con anterioridad de cinco módulos:

- a) Gestor de contenidos.
- b) Módulos de recomendación.
- c) Módulo de gestión de recomendaciones.
- d) Gestión de perfiles.
- e) Gestión de rutas

Este sistema no sólo contiene los módulos indicados anteriormente, sino que implementa varios flujos de información para lograr la interacción requerida por el turista. En primer lugar, el sistema obtiene información del contexto del turista, para que de esta forma se realicen en el servidor los cálculos adecuados y así recomendar los puntos de interés. Una vez recibida esta información, la interacción del turista aceptando, rechazando o valorando las recomendaciones recibidas son transferidas al sistema para mejorar posteriores recomendaciones.

5.3.6. Sistema de contexto

Este módulo realiza un pre-filtrado inicial que eliminará todos aquellos puntos que no sean compatibles con los atributos contextuales. Por ejemplo, si está lloviendo no deben recomendarse actividades en espacios abiertos, como un parque acuático.

Para ello, el sistema de posición y el usuario suministrarán información contextual como lugar, fecha, hora, clima, etc. No obstante, existe otra información contextual importante que debe indicar explícitamente el turista, como por ejemplo si va acompañado de familia, pareja, amigos, etc.

Este sistema utilizará técnicas basadas en análisis formal de conceptos, que permitirá reducir el número de puntos turísticos sobre los que se calculará la predicción de interés (Leiva et al., 2013a). En el siguiente capítulo se describe con detalle el funcionamiento de este sistema.

5.3.7. Sistema de gestión de perfiles

Este sistema tiene como objetivo gestionar información sobre gustos y preferencias de los turistas y las visitas realizadas, permitiéndole acceder además a estos contenidos, pudiendo añadir nuevos, modificarlos o simplemente consultarlos.

Además, debido a que lo habitual es realizar visitas en grupo, este sistema debe proporcionar mecanismos para gestionar simultáneamente perfiles de varios turistas. Así, una de las aportaciones de la solución propuesta es el concepto de turista virtual, que permite al sistema

generar recomendaciones adecuadas a las preferencias de todos los miembros de un grupo de turistas.

5.3.8. Módulos de recomendación

Este sistema está compuesto por los diferentes motores de recomendación, obteniendo de cada uno de ellos una lista de puntos de interés en consonancia con el contexto y de acuerdo a diferentes criterios. En el capítulo 6 se describirá con detalle cada uno de estos motores de recomendación, que son el principal tema de investigación del presente trabajo.

5.3.9. Módulo gestor de recomendación

Se encarga de interactuar con el sistema de recomendación para recoger las recomendaciones generadas por los diferentes motores de recomendación y ordenar los puntos según criterios de preferencia, teniendo en cuenta diferentes parámetros basados en técnicas de hibridación (véase sección 3.4). En un entorno turístico planteamos la técnica en cascada como la más adecuada, haciendo que se reduzca el número de puntos en cada motor de recomendación. En el anexo II se describen los mecanismos de configuración utilizados en el prototipo propuesto.

5.3.10. Sistema gestor de rutas

Partiendo de una serie de puntos recomendados al turista según sus preferencias, este sistema generará rutas de forma dinámica combinando los puntos de interés recomendados al turista con información sobre el orden de visita más frecuente, es decir, el que siguen la mayoría de los turistas y que tiene una mayor aceptación por diversos motivos

(cercanía, precio,...). Por ejemplo, si después de visitar el punto de interés “i”, la mayoría de los turistas visitan el punto de interés “j” y ambos puntos han sido recomendados al turista, parece razonable pensar que una vez un nuevo turista visite el punto “i” se le sugiera visitar el punto “j” y no otro, ya que deben existir razones para que la mayoría decida este orden (cercanía, facilidad para llegar,...). En el capítulo próximo se describe con detalle el procedimiento que se ha diseñado.

5.3.11. Sistema gestor de contenidos

En este sistema se almacena información completa de todos los puntos de interés, sus características, valoraciones, descriptores turísticos-demográficos⁵⁸, atributos contextuales⁵⁹ que pueden afectar una visita, etc. Esta información estará integrada en el gestor de contenidos de la organización de gestión del destino. En las siguientes secciones se describirá con detalle los diferentes mecanismos para almacenar información de los puntos de interés, visitas, valoraciones, etc.

Además, el gestor de contenidos debe almacenar información complementaria sobre los distintos puntos de interés: textos, imágenes y vídeos. Esta gestión debe realizarse en el servidor del destino turístico permitiendo el acceso a la información desde el dispositivo móvil.

⁵⁸ Atributos que describen características socio-demográficas y sobre el tipo de turista en el que se clasifica un usuario.

⁵⁹ Características de un punto de interés relacionadas con el contexto, y que permiten decidir sobre la adecuación de una visita. Así, por ejemplo, la visita a un jardín pierde mucho de su interés si está lloviendo.

5.3.12. Sistema de trazabilidad

Se trata de un sistema que debe registrar las interacciones del usuario con el sistema y los lugares que visita, en qué orden, así como el tiempo que dedica a cada actividad, etc. En el ámbito de este trabajo planteamos respecto a la trazabilidad las dos funciones siguientes:

- a) Gestión de información subjetiva: registro de las valoraciones de puntos de interés realizadas por el turista, lo que permitirá el la actualización de su perfil de cara a perfeccionar las recomendaciones.
- b) Gestión de información objetiva: registro de los puntos de interés visitados, tiempo dedicado y orden de las visitas, etc.

5.4. Ontología RAMCAT

En el capítulo anterior se ha analizado la importancia de las ontologías, así como enumerado un conjunto de ellas en el ámbito turístico. La adecuación de este tipo de tecnología para catalogar información relacionada es cada día más evidente. La creación de ontologías en el área turística facilita la definición formal de diferentes conceptos, que permiten garantizar una representación formal que puede ser compartida y utilizada por cualquier sistema de manera automática (Jakkilinki et al., 2007; Gago et al., 2009; Siricharoen, 2010).

Actualmente, existen diferentes ontologías turísticas (públicas y privadas) que permiten representar tanto aspectos genéricos del ámbito turístico como subdominios más específicos.

En nuestra propuesta, se vio la necesidad de disponer de una ontología que nos permitiera categorizar adecuadamente los recursos para nuestros propósitos de recomendación. Aunque la creación de la misma no es el aspecto principal en nuestro trabajo de investigación, se ha considerado necesario la creación de una ontología propia que se adapte a nuestras necesidades.

En la presente sección se presentan diferentes aspectos del diseño de nuestra ontología. El desarrollo de una ontología se puede realizar utilizando diferentes lenguajes, como RDF, DAML u OWL. Nos hemos decantado por utilizar este último como lenguaje de implementación, ya que permite formalizar tanto conceptos como relaciones. La elección de este lenguaje se debe fundamentalmente a que es el más utilizado actualmente y además su editor (Protégé) es de código abierto.

La ontología que se ha desarrollado, obviamente puede ser modificada por los agentes del destino turístico, que pueden incluir nuevas instancias a medida que estén disponibles y cuando se considere necesario.

En concreto, los atributos formalizados en nuestra conceptualización para los puntos de interés turístico son:

- a) TURID: código que permite identificar de forma única cada punto de interés turístico.
- b) Nombre o título del punto de interés turístico.
- c) Palabras clave: describen todas las características y subcategorías en la que puede catalogarse el punto turístico.
- d) Descripción: breve resumen del punto de interés turístico.

- e) Categoría: atributo principal que describe al punto de interés (categoría de mayor nivel). Por ejemplo la categoría Restaurante tiene un nivel mayor que Restaurante Mexicano.
- f) Horario: informa de los días y del horario en los que es posible la visita.
- g) Precio: información respecto al coste de visita del punto de interés.
- h) Minusvalía: información especial para turistas discapacitados (accesibilidad, etc.).
- i) Localización: geolocalización del punto de interés.
- j) Tiempo: duración media de la actividad turística.
- k) Grupo: indica las características demográficas específicas de los visitantes potenciales del punto de interés en función de su edad, profesión, nivel cultural, nivel económico, etc. Por ejemplo, un parque acuático está más orientado a niños y acompañantes jóvenes que a personas jubiladas.
- l) Propiedad: indica si el punto de interés pertenece a una entidad pública o privada.
- m) Clima: señala las características climáticas deseables para realizar la visita al punto de interés de forma satisfactoria.
- n) Relación: este atributo informa de la existencia de algún tipo de relación con otro punto de interés. Por ejemplo un museo puede estar dentro de una catedral.
- o) Idioma: idiomas en los que existe información en ese punto de interés.

p) Características del producto: se trata de una lista con características concretas de ese punto de interés. Por ejemplo, en el caso de un restaurante pueden ser atributos relacionados con el tipo de comida, si dispone de terraza, etc. Cada una de esas características es una tupla formada por nombre de característica y valor. Así, una característica “económico” valdrá 1 si se considera muy económico, 0,5 si se considera un precio medio y un 0 nada económico. Estos valores son indicados por expertos del destino. En la tabla 5.1 podemos observar como para una serie de puntos de interés, un especialista del destino describe, asignando un valor entre 0 y 1, características relativas al tipo de espacio.

	Diseño	
	Espacio abierto	Espacio cerrado
i_1	0.3	0.8
i_2	0.1	0.8
i_3	0.3	0.8
i_4	0.3	0.8
i_5	0.1	0.9
i_6	0.9	0.2

Tabla 5.1. Descripción de características de tipo de espacio

q) Características del turista: se trata de una lista que describe el grado de adecuación de ese POI asociado a cada tipo de turista (descriptores turísticos). Por ejemplo, el nacimiento de un río tendrá asociada una tupla formada por tipo de turista y un valor

entre 0 y 1, el cual indicará el grado de adecuación de ese punto de interés a cada tipo de turista. Así, para un turista cultural tendrá un valor 0,5 y para un turista de naturaleza un valor 1.

Para cada atributo descrito anteriormente, se define un conjunto de posibles valores que proporcionan descripciones precisas sobre los diferentes puntos de interés. La ontología RAMCAT OWL ha sido diseñada usando clases, instancias y propiedades OWL. Así, cada punto de interés será identificado de forma única por el valor de su atributo TURID, que identifica a cada instancia de la ontología. Cada una de las instancias pertenece a clases que estarán categorizadas jerárquicamente. Para definir la jerarquía se ha tomado como base el tesauro de la OMT y así como varias ontologías turísticas (Euzénat, Remize y Ochanine, 2003; Zou et al., 2003; Siorpaes y Bachlechner, 2006; Ou et al., 2008; Barta et al., 2009; Minguez, Berrueta y Polo, 2009; Siricharoen, 2010; Herrero et al., 2012). (descritas brevemente en el capítulo anterior), en las que es posible identificar los distintos tipos de actividades o puntos de interés. Este conjunto de atributos define adecuadamente las características semánticas de los distintos puntos. En el siguiente apartado se describe la jerarquía detalladamente.

5.4.1. Jerarquía de actividades turísticas

La especificación del tesauro de la OMT establece una clasificación detallada de términos específicos del turismo (OMT, 2001). Se trata de un instrumento destinado a indexar conceptos turísticos para su utilización en el diseño de sistemas informáticos, pretendiendo estandarizar y normalizar mediante un lenguaje común internacional,

con el objetivo de mejorar la promoción e información en el sector turístico.

El conjunto de términos asociados al turismo se reparte en campos semánticos que representan la actividad turística. Para el sistema propuesto, se han seleccionado los campos aplicables a nuestro modelo, modificando y añadiendo algunos según los objetivos de nuestros métodos de recomendación.

En la tabla 5.2 pueden observarse las diferentes categorías genéricas que hemos utilizado y que nos permitirán identificar otras más específicas. A partir de esta estructura (que consideramos la más apropiada para nuestros propósitos y a nuestro ámbito), la construcción de la jerarquía de actividades turísticas de nuestra ontología es inmediata.

Categoría	
1	Actividades deportivas
2	Ecología
3	Economía
4	Instalaciones turísticas
5	Alojamiento
6	Ocio
7	Eventos turísticos
8	Monumentos y patrimonio
9	Servicios turísticos
10	Transporte

Tabla 5.2. Clasificación de categorías turísticas.

Cada una de estas categorías estará identificada por un código, para que resulte más sencillo inferir la especialización. Así, si la categoría Monumentos y patrimonio está identificada mediante el código 8, se puede identificar con los códigos 8.1, 8.2... a diferentes especializaciones de la misma, para referirnos a subcategorías de la categoría principal.

En la tabla 5.3 podemos observar algunas de las subcategorías de la categoría Monumentos y Patrimonio. Los diferentes puntos de interés pertenecen siempre a las clases hoja de la jerarquía. De esta forma, el sistema podrá inferir el nivel de interés de un turista en relación a la categoría a la que pertenece un punto de interés y también en relación a las categorías que le anteceden en la jerarquía definida.

CATEGORÍA	CÓDIGO
Monumentos	8
Monumento Civil	8.1
Auditorio	8.1.1
Biblioteca	8.1.2
.....
Monumento Industrial	8.2
Bodega	8.2.1.
Depósito	8.2.2.
Fábrica	8.2.3.
Lonja	8.2.4
.....

Tabla 5.3. Clasificación de subcategorías de la categoría Monumentos y Patrimonio

Además, cada punto de interés puede pertenecer a una o varias categorías, ya que en otro caso, la ausencia de clasificación múltiple haría perder precisión en la representación del conocimiento del punto turístico.

5.4.2. Diseño de la ontología

Como hemos indicado anteriormente, la herramienta utilizada para diseñar la ontología es Protégé, en la cual se distinguen clases, propiedades e instancias. Las instancias representan los objetos dentro del dominio de interés, las propiedades son las relaciones entre las instancias y las clases son los elementos que contienen las instancias. En el apéndice 1, se detallan los aspectos de la ontología propuesta que ha sido necesario crear para poder desarrollar nuestros modelos de recomendación.

5.5. Definiciones preliminares

En esta sección se define un conjunto de conceptos necesarios para describir adecuadamente el mecanismo de representación del conocimiento de los puntos de interés y de los turistas, así como de los motores de recomendación.

Definición 5.1: Se denomina objetivo de un turista a una petición de recomendación de visita a puntos turísticos o ruta de puntos turísticos, los cuales se deben encontrarse a una distancia menor o igual a una dada (y configurable) del lugar en el que se encuentra el turista o se encontrará en un momento determinado.

El turista representará su objetivo como una búsqueda, de acuerdo a uno o varios de los siguientes criterios:

- a) Segundo las preferencias del perfil del turista.
- b) Especificando categorías positivas o negativas, es decir, indicando qué tipo (características) de puntos desea visitar o qué tipo de puntos no desea visitar.
- c) Indicando la información contextual explícitamente.

Definición 5.2: Denominamos sesión al periodo de tiempo que transcurre entre una conexión del usuario al sistema y la posterior desconexión. Durante una sesión el turista puede realizar varias búsquedas mediante la definición de diversos objetivos.

Definición 5.3: Se denomina acción a toda operación que realice el turista con el sistema de realidad aumentada aplicada al turismo basada en el contexto. Existen diferentes tipos de acciones, como pueden ser búsqueda de puntos de interés, petición de recomendación de ruta, evaluación de un punto de interés, selección o desección de puntos de interés recomendado, etc. Todas estas acciones pueden implicar cambios en el perfil que el sistema tenga del turista.

Definición 5.4: Se denomina valoración a la acción mediante la cual el usuario valora un punto de interés.

La valoración o evaluación es una acción fundamental, pues consigue retroalimentar el sistema para definir más fielmente el perfil de preferencias del usuario. Existen dos tipos de valoraciones: general (el turista evalúa la visita con un único valor) y detallada (el turista evalúa

la visita en diferentes categorías, atributos y según parámetros contextuales).

A continuación, vamos a indicar diferentes notaciones que utilizaremos en posteriores definiciones:

- a) Notaremos U al conjunto de usuarios. En nuestro caso será el conjunto de turistas que utilizará el sistema de RA. Los usuarios los notaremos como u_1, u_2, \dots .
- b) Notaremos G a una categoría o estereotipo de usuarios. Cada usuario pertenece a una categoría o estereotipo de usuarios. Un usuario u es asignado a una categoría según su afinidad a las preferencias predeterminadas de ese grupo, así como a características demográficas.
- c) Notaremos I al conjunto de posibles ítems a recomendar. En el caso que nos ocupa serán los diferentes puntos de interés o actividades turísticas gestionadas por el destino. Los ítems individuales serán denotados como i_1, i_2, \dots .

Definición 5.5: Sea C el conjunto de categorías de ítems. Cada ítem $i \in I$, estará asignado a una o varias categorías, es decir, si $C = \{c_1, c_2, \dots, c_n\}$, donde c_i es una categoría (por ejemplo, museo, playa, etc.), diremos entonces que todo punto de interés i debe pertenecer al menos a un $c_i / c_i \in C$. Estas categorías pertenecen a la ontología definida. En la sección 5.6 precisaremos la jerarquía definida para las categorías de ítems.

Definición 5.6: El nivel de pertenencia de un ítem $i \in I$ o subcategoría $c' \in C$ a una categoría $c \in C$ (denotado k_c^i o k) es un valor entre 0 y 1.

El valor 1 indica un nivel de pertenencia máximo y el valor 0 un valor mínimo. Para ello se definen dos funciones de pertenencia:

$$F_{pertenencia\ item}: I \times C \rightarrow [0,1] / F_{pertenencia}(i, c) = k$$

$$F_{pertenencia\ categoria}: C \times C \rightarrow [0,1] / F_{pertenencia}(c', c) = k$$

Definición 5.7: El contexto (denotado T) se define como la situación en la que se encuentra el turista que solicita una recomendación y se describe mediante varios atributos que se tendrán en cuenta en la recomendación (compañía, posición, fecha, hora, clima, duración actividad, precio...). La valoración de estos atributos es opcional, y en algunos casos es obtenida por el dispositivo (como la fecha y hora, por ejemplo), mientras que en otros (como la compañía) la valoración debe ser indicada explícitamente por el turista.

Definición 5.8. Valoración del ítem i del usuario u (v_u^i). Cada valoración puede realizarse en presencia y ausencia de diferentes características contextuales. Las valoraciones pueden realizarse asignando un valor numérico o mediante variables lingüísticas, es decir, pueden estar en diferentes escalas, lo que hará necesario el definir operaciones para poder convertir las calificaciones a una única escala numérica.

Definición 5.9: Valoración general que tiene el ítem i en la categoría c (r_i^c). Un usuario cuando realiza una valoración puede realizarla de forma general (dando un único valor) o de forma pormenorizada (indicando diferentes valoraciones por categorías, características, etc.) Se hará necesario en este último caso definir operaciones para obtener la valoración general a partir de la valoración pormenorizada. La forma de calcularla será:

$$r_i^c = \frac{\sum_{m=1}^n v_{u_m}^i}{n}$$

donde $v_{u_m}^i$ será la valoración dada por el usuario a ese ítem⁶⁰. Como se ha descrito anteriormente esta valoración puede ser general o por categoría si el turista realiza la valoración de forma pormenorizada. Al final de la sección 5.6.1 mostramos un ejemplo que ilustra la forma de calcularlo.

Definición 5.10: Influencia del atributo contextual t_j en la categoría c en el ítem i (r_{itj}^c). Es un valor que indica cómo influye un atributo contextual t_j en la valoración. Por ejemplo, una categoría determinada puede tener mayor interés dependiendo de la compañía con la que se realiza la visita, así un parque de atracciones tiene una influencia máxima si va acompañado de niños, pero su interés puede ser menor si va sólo con su pareja. La forma de calcularlo sería en aquellos atributos que puedan tener influencia en la valoración:

$$r_{itj}^c = \frac{\sum_{m=1}^n v_{u_m}^i}{\frac{n}{r_i^c}}$$

donde $v_{u_m}^i$ es la valoración que ha realizado el turista u_m del punto i , estando presente el contexto t_j . En los atributos contextuales donde no exista influencia en las calificaciones el valor por defecto será uno. En la sección 5.6.2 se ilustrará con un ejemplo la forma de calcularlo.

⁶⁰ Por sencillez en la expresión y para una mejor comprensión, supondremos las valoraciones dadas en forma numérica, aunque debido a que se utilizarán escalas lingüísticas serán necesarias definir operaciones para trasladar las valoraciones de una escala a otra.

Definición 5.11: Sea V_i^c la valoración del ítem i en la categoría c , definida como un vector que tendrá $2 \cdot n + 1$ elementos, siendo n el número de características contextuales.

$$V_i^c = (r_i^c, r_{it1}^c, r_{it2}^c \dots r_{itn}^c, r_{it\bar{1}}^c, r_{it\bar{2}}^c \dots r_{it\bar{n}}^c)$$

donde r_i^c es la valoración general del ítem i en la categoría c (definición 5.9) y r_{itj}^c , es el valor del atributo contextual j para el ítem i en la categoría c (definición 5.10). Para los atributos contextuales, denominamos r_{itk}^c al factor de influencia del atributo contextual k para el ítem i asignado a la categoría c sobre ese punto turístico. Por otro lado, $r_{it\bar{k}}^c$ es el factor de influencia cuando no se da el atributo contextual k para el ítem i asignado a la categoría c .

Definición 5.12: Valoración general que tiene el usuario u en la categoría c (r_u^c).

Este valor se puede obtener inicialmente en el proceso de registro del usuario u , para posteriormente según se vayan valorando (v_u^i) ítems pertenecientes a esa categoría ir variando. En el caso de no tener este valor inicialmente, se obtendrá exclusivamente de las visitas que realice el turista.

Definición 5.13: Influencia del atributo contextual m en la categoría c por el usuario u (r_{um}^c). Es un valor entre 0 y 1 que indica cómo influye un atributo m en la valoración.

Por defecto los valores r_{um}^c tomarán valor unidad y sólo en aquellos casos donde esa categoría influya en un atributo negativamente, tomará un valor entre 0 y 1. Por ejemplo, si no desea que le recomienden restaurantes románticos cuando va con niños, ese atributo contextual

tomará un valor nulo en esa categoría. Este valor unitario irá disminuyendo cuando la valoración de un punto recomendado sea inferior a lo esperado, si se ha presentado ese contexto.

Definición 5.14: Valoración de la categoría c por el usuario u (V_u^c). Se trata de un vector de dimensión $k+1$, donde k es el número de atributos contextuales asociados a la categoría. Así, su estructura será

$$V_u^c = (r_u^c, r_{u1}^c, r_{u2}^c, \dots, r_{uk}^c),$$

donde r_u^c es la valoración general de esa categoría, mientras que r_{uk}^c es la valoración que realiza el usuario u en la categoría c cuando se presenta el atributo contextual k . Es importante destacar como indicamos en la definición 5.9, que el turista puede realizar valoraciones generales o pormenorizadas.

Definición 5.15: Se define descriptor como un atributo o característica que puede definir a un turista o al tipo de turistas que visitan un determinado ítem. Denotaremos D al conjunto de descriptores turísticos y demográficos que indica la tipología de un turista.

Un ejemplo de descriptor turístico puede ser turismo de naturaleza, turismo cultural, etc. Un ejemplo de descriptor demográfico podría ser el nivel de estudios. Cada destino puede crear los descriptores oportunos. Cada turista y cada recurso turístico podrán definirse mediante valores para todos los descriptores que se consideren adecuados.

Definición 5.16: Sea d_k^u la valoración del descriptor k para el usuario u .

Definición 5.17: Sea d_k^i la valoración del descriptor k para el punto turísticos i .

Definición 5.18: Se denomina preferencia de recomendación específica del usuario u (SRP^u) a la estructura que almacena el conjunto de todos los valores V_u^c del usuario u para todas las categorías. Se trata de un vector formado por tantos elementos como categoría valoradas por el usuario u .

Definición 5.19: Cada turista u debe estar categorizado dentro de un estereotipo de turista por semejanza, creándose una estructura denominada SGP^u (Preferencias de especificación de grupo), que almacena las preferencias y características del estereotipo mediante descriptores demográfico-turísticos.

Definición 5.20: Preferencias de especificación históricas SHP^u , que almacenará información sobre las visitas realizadas por el turista u . En este caso la estructura utilizada es un vector que tiene tantos elementos como visitas realizadas por el turista.

Definición 5.21: Preferencias específicas de planificación SPP^u , que almacenará información sobre las características de los productos que desea visitar, así como su tipo. Se trata de un vector formado por tuplas de tres elementos: el primero es la categoría, el segundo es una característica y el tercero el valor deseable de esa característica por parte del turista. Por ejemplo, un turista puede desear ir a un restaurante que sea bastante económico.

Se dirán que dos turistas tienen en una categoría gustos idénticos, similares o distintos, según se encuentren en una de las siguientes situaciones⁶¹:

Definición 5.22: Dos turistas u_1 y u_2 tienen gustos idénticos si se cumple que tienen la misma valoración para todas las categorías⁶².

$$\forall c \in C, V_{u_1}^c = V_{u_2}^c$$

Definición 5.23: Dos turistas u_1 y u_2 tienen gustos similares si tienen todas las categorías valoradas de idéntica manera, excepto algunas que tendrán una diferencia de valoración inferior a un umbral definido en el sistema⁶³.

$$\forall c \in C, |V_{u_1}^c - V_{u_2}^c| \leq \text{umbral}$$

Definición 5.24: Dos turistas u_1 y u_2 tienen gustos diferentes si tienen alguna categoría con una diferencia de valoración mayor a un umbral definido en el sistema⁶⁴.

$$\exists c \in C, |V_{u_2}^c - V_{u_1}^c| > \text{umbral}$$

5.6. Representación del conocimiento de los puntos turísticos

A continuación, se describe cómo representar el conocimiento de los diferentes puntos de interés turístico. El subsistema de gestión de la

⁶¹ Para facilitar la lectura y comprensión, supondremos que las valoraciones siempre tienen una correspondencia numérica. Es muchas ocasiones utilizaremos valoraciones lingüísticas, pero el proceso de conversión de valores lingüísticos a numéricos, se abordará formalmente en los modelos de recomendación descritos en el capítulo 6.

⁶² No confundir la notación V_i^c con V_u^c . La primera es valoración de un ítem en una categoría, y la segunda es valoración de un usuario a una categoría.

⁶³ En la tabla 5.6 se ha expuesto una aplicación de esta definición suponiendo que se valora una categoría en una escala del 1 al 5, donde el umbral es 1.

⁶⁴ Igual que la nota anterior

información utilizará la ontología RAMCAT para poder representar y manipular la información de las preferencias del turista, así como para catalogar los diferentes puntos a recomendar.

Además de usar las propiedades definidas en la ontología para categorizar los diferentes puntos de interés turístico, es necesario que el sistema utilice otras propiedades de la ontología descrita en la sección 5.4 (por ejemplo la historia de un monumento, tiempo estimado de una visita, horario de apertura, precios medios, etc.). Esta información será de mucha utilidad al turista cuando, por ejemplo, solicite una recomendación de ruta, calculando tiempos estimados de desplazamiento, costes, etc.

En la figura 5.4 podemos observar parte de la ontología creada con información sobre diferentes tipos de destinos y puntos de interés.

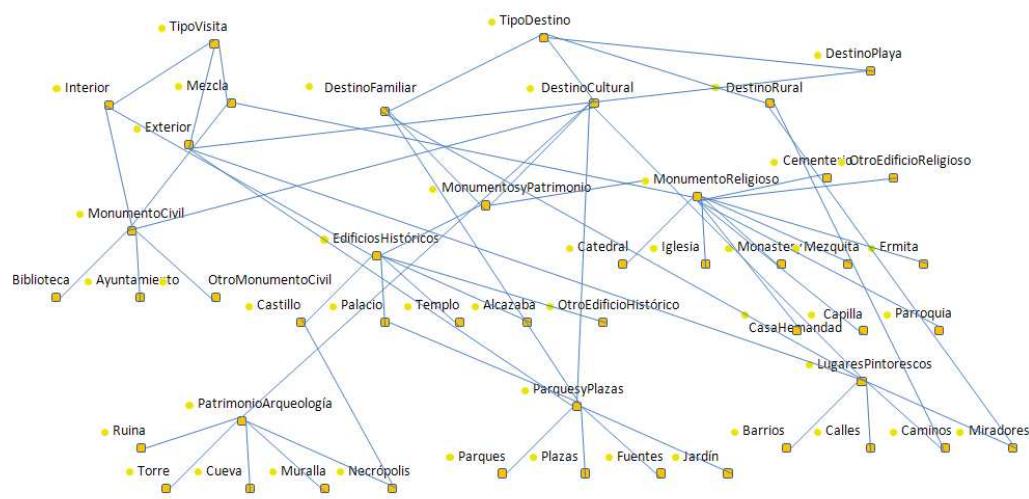


Figura 5.4. Visión parcial de la ontología desarrollada para RAMCAT

La ontología consta de 377 conceptos que son utilizados para realizar las recomendaciones. Cada punto turístico está etiquetado con uno o varios conceptos de la ontología, de forma que las diferentes relaciones

existentes entre los nodos permitirán relacionar las preferencias del turista sobre un determinado punto turístico o actividad. Por ejemplo, si un turista visita un monasterio, el sistema deduce que tiene interés en visitar edificios religiosos, pero también edificios históricos, ya que los monasterios están asociados a ambos conceptos.

Aunque se describirá detalladamente en posteriores secciones, es importante destacar que el sistema propuesto utiliza técnicas para enriquecer el perfil del turista, actualizando dinámicamente sus preferencias a partir de sus interacciones y visitas. Así, el hecho de que un turista seleccione la opción de mostrar información detallada de un determinado punto turístico o elimine una actividad recomendada para una ruta, permite aumentar o actualizar la información existente del propio turista. Toda esta información, junto con las valoraciones que realicen los turistas, deberá ser trasladada a la estructura que almacena sus preferencias y a la estructura de conocimiento de los puntos turísticos.

5.6.1. Taxonomía propuesta

El sistema de gestión de la información de ítems utiliza una taxonomía a partir de la ontología desarrollada, tanto para representar y manipular la información de las preferencias del turista como para catalogar los diferentes ítems.

Las entidades en una taxonomía se organizan en una estructura jerárquica conectadas por una relación “es una”, donde los niveles de clasificación más específicos corresponderán a los niveles inferiores. Las entidades representan conceptos comunes de un dominio del turismo; en

cambio, los niveles más específicos de la taxonomía serán las actividades o ítems.

La taxonomía puede definirse como un grafo dirigido ponderado (V, E), donde V es el conjunto de nodos del grafo, los cuales representan conceptos de la ontología descritos como categorías, subcategorías de ítems y los diferentes ítems (que en este trabajo serán los elementos susceptibles de ser recomendados). El conjunto de nodos no hoja⁶⁵ representan categorías, mientras que los nodos hoja del grafo serán los ítems. Por último, E es el conjunto de relaciones entre nodos del conjunto V . Cada arco tiene asociado una función de ponderación que se define de la siguiente forma:

$$W:E \rightarrow [0,1]$$

A continuación definimos los distintos tipos de conectores (Figura 5.5):

- Conexión categoría-categoría, que vincula una categoría c_i con una subcategoría c_j . Este tipo de flechas sólo conecta categorías con subcategorías exclusivas, es decir, que no es compartida por ninguna otra, por lo que tiene por defecto un valor de pertenencia unitario. En otros términos, si e_{ij} es la relación que une a las categorías c_i y c_j y no existe otra categoría c_k de la cual sea subcategoría c_j , entonces $W(e_{ij}) = 1$.
- Conexión categoría-categoría ponderada, que vincula una categoría c_i con una subcategoría c_j teniendo una ponderación k (véase la definición 5.6) basada en un grado de pertenencia con un valor

⁶⁵ Un nodo que no tiene hijos se conoce como nodo hoja.

entre [0,1]. Este tipo de conexión une categorías con subcategorías compartidas. En este caso la flecha va acompañada de un valor k que indica el grado de pertenencia a dicho concepto, indicando el 1 máxima pertenencia y el 0 mínima pertenencia.

- Conexión categoría-ítem, que vincula una categoría c_j con un punto de interés turístico i , teniendo una ponderación basada en un grado de pertenencia con un valor entre [0,1], tal como definimos anteriormente.

Cada actividad puede tener asociadas valoraciones según diferentes atributos de contexto. Estos valores pueden ir variando con el tiempo, según las visitas y calificaciones de los turistas que usan el sistema. El valor general de un punto i en la categoría c , lo hemos denotado r_i^c (definición 5.9). Dado que este valor puede alterarse por diferentes parámetros contextuales, necesitamos una estructura que almacene los diferentes valores. A esta estructura la denominamos V_i^c (definición 5.11), que representa la valoración de ese punto en la categoría c y cómo le influyen los diferentes parámetros de contexto. En este caso, el valor k (definición 5.6) indica el grado de pertenencia del punto de interés a esa categoría (valor entre 0 y 1, aunque normalmente suele llevar el valor unidad o valores próximos). A las categorías que se vinculan directamente con un punto de interés turístico se le denomina categoría hoja o terminal.

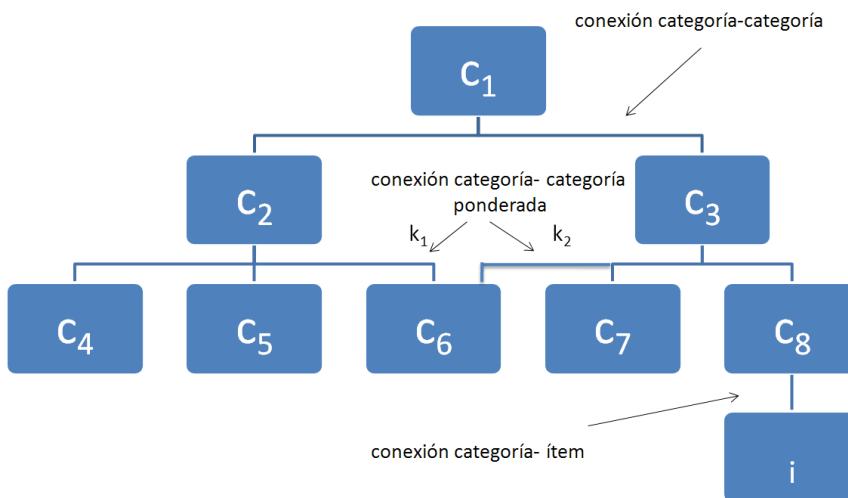


Figura 5.5. Tipos de conexiones en la taxonomía propuesta

Así, en el nodo de la taxonomía correspondiente a Destino Cultural, encontramos categorías como por ejemplo Edificios, que a su vez se subdivide en Edificios religiosos, Edificios históricos y así sucesivamente hasta llegar a los nodos hoja que serán los diferentes puntos de interés del destino. Como se ha indicado, los nodos hoja pueden pertenecer a una o varias categorías, como se puede observar en la figura 5.6, en la que el punto de interés Catedral de Málaga está relacionado con tres categorías: catedrales, jardines y pintura.

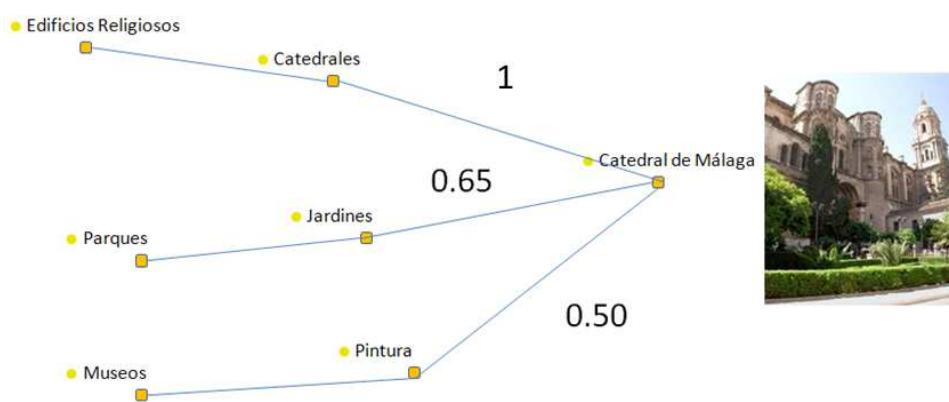


Figura 5.6. Visión parcial de la taxonomía correspondiente al punto de interés

Catedral de Málaga

Así, el valor que aparece en la línea que conecta cada concepto con el punto de interés indica el grado de pertenencia del punto en relación a cada concepto. Así, el punto Catedral de Málaga está asociado a un valor de pertenencia del 100% (valor 1) al concepto “Catedrales”. El mismo punto también posee una relación con el concepto “Jardines”, con un valor de pertenencia del 65% (valor 0,65) y con el concepto “Pintura”, cuyo valor de pertenencia es del 50% (valor 0,5).

Cada nodo hoja representa un ítem que tiene asignado un valor V_i^c por cada categoría (definición 5.11) con la que está relacionado, valor que se irá retroalimentando con las evaluaciones que realizan los turistas. El turista podrá evaluar la actividad de forma general o teniendo en cuenta los atributos contextuales. Para ello utilizará una escala lingüística que posteriormente le asignará una correspondencia numérica el sistema⁶⁶. Supongamos el siguiente ejemplo de escala lingüística:

- Punto muy interesante: puntos que deben visitarse sin duda por los turistas que tengan un perfil muy relacionado con esta categoría. En una escala numérica podríamos hacerle corresponder el valor máximo, por ejemplo un diez en una escala [0,10].
- Punto bastante interesante: deben visitarse siempre que se pueda. Siguiendo el ejemplo anterior podría considerarse un valor 8 en una escala [0,10].

⁶⁶ Para facilitar la comprensión asimilaremos la escala lingüística a una escala numérica. En el capítulo seis se describirá la utilización de técnicas difusas para el manejo adecuado de las diferentes escalas lingüísticas.

- Punto con interés normal: se trata de puntos de interés que no tienen demasiada relevancia. En este caso se podría considerar el valor 5 o 6 en una escala [0,10]
- Punto poco interesante: puntos que tienen un interés relativo. En una escala [0,10] podríamos asignarle un valor 4.
- Punto con interés bajo: puntos que tienen muy poca relevancia en esa categoría. Continuando con el mismo ejemplo, podría asignarse un valor nulo o 1 o 2.

A continuación, vamos a ilustrar mediante un ejemplo cómo se almacenan las valoraciones de los turistas que visitan la Catedral de Málaga. Para simplificarlo, supondremos un número de doce valoraciones de diferentes usuarios (u_1 hasta u_{12}), el número de categorías en las que se puede recomendar la Catedral del Málaga son tres (A – Categoría Catedrales; B – Categoría Jardines; C – Categoría Pintura) y tres atributos contextuales⁶⁷ que pueden influir en la valoración final de la visita: C_1 (Visita con niños), C_2 (Visita en Festivo) y C_3 (Clima soleado).

Asumiremos que el valor -1 en un atributo contextual representa que no tiene influencia en las valoraciones, mientras que el valor 1 significa que es cierta la presencia de esa situación de contexto y el valor 0 indicará la ausencia de la misma. Así, por ejemplo, visitar una catedral un día soleado o lluvioso no debe influir en las valoraciones, pues es una actividad que se realiza en un sitio cerrado.

⁶⁷ La inclusión en este ejemplo de atributos contextuales se debe a que en la sección siguiente dedicada a éstos se empleará el mismo ejemplo.

Por simplicidad, en el ejemplo supondremos que cada usuario valora el ítem únicamente en una categoría. No obstante, el sistema debe permitir que cada usuario pueda valorar el ítem en varias categorías e incluso por características, siendo en estos casos una evaluación exhaustiva. La columna categoría en la tabla 5.4 indicará la categoría por la que realiza la valoración el turista, que supondremos que es la categoría por la que se le ha recomendado dicha actividad.

Supongamos las siguientes valoraciones en una escala de 0 a 10:

Usuario (u_m)	Valoración ($v_{u_m}^i$)	Categoría	Atributos contextuales		
			C ₁	C ₂	C ₃
u ₁	8	A	1	1	-1
u ₂	10	A	0	1	-1
u ₃	6	C	0	0	-1
u ₄	8	A	1	1	-1
u ₅	8	A	0	0	-1
u ₆	8	B	0	0	1
u ₇	10	C	0	1	-1
u ₈	4	B	0	1	0
u ₉	6	A	1	1	-1
u ₁₀	10	A	0	0	-1
u ₁₁	10	A	0	1	-1
u ₁₂	6	C	1	1	-1

Tabla 5.4. Ejemplo de valoraciones en la Catedral de Málaga

A partir de estos datos podemos obtener los siguientes resultados r_i^c (definición 5.9), para el punto turístico Catedral de Málaga⁶⁸, sabiendo que:

$$r_i^c = \frac{\sum_{m=1}^n v_{u_m}^i}{n}$$

Donde m es el número de turistas que ha valorado la actividad i en la categoría c , y $v_{u_m}^i$ es la valoración dada por el turista en esa categoría, quedando entonces los siguientes resultados en el ejemplo:

$$r_{\text{Catedral de Málaga}}^{\text{Categoría Catedrales}} = \frac{8 + 10 + 8 + 8 + 6 + 10 + 10}{7} = 8.57$$

$$r_{\text{Catedral de Málaga}}^{\text{Categoría Jardines}} = \frac{8 + 4}{2} = 6.00$$

$$r_{\text{Catedral de Málaga}}^{\text{Categoría Pintura}} = \frac{6 + 10 + 6}{3} = 7.33$$

Al poder pertenecer un ítem a varias categorías, tendrá por tanto calificaciones r_i^c diferentes, que irán actualizándose con las valoraciones que hagan nuevos turistas.

5.6.2. Atributos contextuales

En esta sección vamos a ilustrar la forma de calcular la influencia de un atributo contextual t en la categoría c en el ítem i , r_{it}^c (definición 5.10).

Como hemos indicado con anterioridad, cuando un atributo o propiedad contextual de un punto turístico no tiene ningún tipo de influencia sobre la valoración general que pueda tener un turista, el

⁶⁸ La forma de calcularse es mediante la media de las valoraciones realizadas por los diferentes usuarios que han valorado esa actividad.

valor que se le asigna es -1; mientras que si una propiedad tiene una influencia sobre la valoración general que pueda tener un turista, el valor que se le asigna es un valor positivo⁶⁹. Más formalmente, se denomina r_{it}^c (véase la definición 5.10) al valor que determina cómo influye el atributo de contexto t en la valoración del punto turístico i en la categoría c . A continuación exponemos un ejemplo.

Supongamos el atributo contextual relacionado con el clima. Para visitar por ejemplo una catedral, este atributo no tiene importancia en el resultado de la valoración de los turistas que realizan la visita, por lo que este atributo contextual tendría asignado un valor -1 en ese punto turístico. Pero en el caso de visitar un parque, el clima sí debe tener importancia. De esta manera, cuando haga mal tiempo, el valor de recomendación de ese punto será menor.

Para el cálculo de la influencia del atributo contextual t en la categoría c en el ítem i , r_{it}^c (definición 5.10) utilizaríamos las siguientes expresiones, donde $v_{u_m}^i$ es la valoración que ha realizado el turista u_m del punto i , estando presente el contexto t y $v_{\bar{u}_m}^i$, que es la valoración que ha realizado el turista u_m del punto i , *no* estando presente el contexto t .

$$r_{it}^c = \frac{\sum_{m=1}^n v_{u_m}^i}{\frac{n}{r_i^c}}$$

de la misma manera se calcularía:

⁶⁹ Tal como hemos indicado anteriormente, para facilitar la comprensión asimilaremos la escala lingüística a una escala numérica. En el capítulo seis se describirá la utilización de técnicas difusas para el manejo adecuado de las diferentes escalas lingüísticas

$$r_{it}^c = \frac{\frac{\sum_{m=1}^n v_{um}^i}{n}}{r_i^c}$$

donde n es el número de valoraciones realizadas por todos los turistas que han visitado ítem i .

Volviendo al ejemplo anterior, para el caso del punto turístico Catedral de Málaga se obtienen los siguientes valores:

$$r_{Catedral\ de\ Málaga-C1}^{Categoría\ Catedrales} = \frac{\frac{8+8+6}{3}}{8.57} = 0.86$$

$$r_{Catedral\ de\ Málaga-C1}^{Categoría\ Catedrales} = \frac{\frac{10+8+10+10}{4}}{8.57} = 1.12$$

$$r_{Catedral\ de\ Málaga-C2}^{Categoría\ Catedrales} = \frac{\frac{8+10+8+6+10}{5}}{8.57} = 0.98$$

$$r_{Catedral\ de\ Málaga-C2}^{Categoría\ Catedrales} = \frac{\frac{8+10}{2}}{8.57} = 1.05$$

$$r_{Catedral\ de\ Málaga-C3}^{Categoría\ Catedrales} = 1$$

$$r_{Catedral\ de\ Málaga-C3}^{Categoría\ Catedrales} = 1$$

El valor uno de los dos últimos valores calculados se debe (véase definición 5.10) a la ausencia de influencia de este atributo contextual en las valoraciones.

5.6.3. Atributos no contextuales

En algunos motores de recomendación el turista puede especificar las características específicas de los productos que desea visitar. Así, por ejemplo un turista puede desear ir a restaurantes que sean económicos, que tenga espacios libres para que puedan jugar los niños, pero que a su vez tenga aire acondicionado en las zonas cerradas.

Por tanto, los puntos de interés tienen una serie de atributos o características que emplearán los motores de recomendación para filtrarlos según las preferencias del turista que solicita la recomendación. A estos atributos los hemos denominado “Características del producto” tal como describimos en la sección 5.4. El número de atributos depende del ítem, pero todos los ítems de una misma categoría tienen atributos comunes. Lo ideal sería que todos los puntos tuvieran el mismo número de atributos no contextuales, pero por ejemplo un atributo servicio puede tener sentido en un restaurante y no en una visita a un museo.

En la tabla 5.5 podemos observar algunas características de algunos puntos. Las valoraciones son suministradas por los expertos del dominio en el destino turístico, asignando un valor en la escala [0,1]. Por ejemplo en la tabla podemos ver que el ítem i_2 es nada barato, pero bastante caro.

	Ambiente			Precio		
	Tranquilo	Animado	Pintoresco	Barato	Moderado	Caro
i_1	0.8	0.5	0.2	0.7	0.3	0.3
i_2	0.9	0.2	0.1	0	0.1	0.9
i_3	0.3	0.8	0.1	0.9	0.3	0.1
i_4	0.2	0.8	0.9	0.9	0.5	0.1
i_5	0.3	0.8	0.7	0.9	0.5	0.3
i_6	0.3	0.8	0.8	0.5	0.7	0.8

Tabla 5.5. Valoración de atributos no contextuales

5.7. Representación del conocimiento de los turistas

En esta sección vamos a describir los mecanismos para la creación de los diferentes perfiles de los turistas que utilizarán el sistema propuesto.

Para ello, se propondrán diferentes técnicas para obtener la información del perfil inicial de preferencias del turista.

Debido a que la propuesta utilizará diferentes técnicas de recomendación, se hace necesario manejar otros tipos de informaciones sobre el turista diferente a sus preferencias. Una información muy interesante del turista es conocer el grupo demográfico-turístico al que pertenece (en nuestro caso lo denominaremos estereotipo, tal como indicamos en la sección 5.5), además de permitir que el turista pueda indicar explícitamente las características de los ítems que quiere visitar (en nuestro caso lo denominaremos necesidades actuales del turista).

5.7.1. Creación de perfiles individuales

Debido a que el sistema propuesto de recomendación es híbrido, se necesita almacenar información muy diversa de los turistas que van a realizar la visita, es decir, de los usuarios del sistema. El sistema posee la capacidad de adaptarse a las necesidades personales de cada usuario y debe generar recomendaciones independientemente de:

- a) La disponibilidad de información de visitas previas. El no disponer de información previa hace que varias técnicas de recomendación no sean adecuadas.
- b) La calidad de la información del perfil.
- c) La existencia de información explícita sobre las características de los ítems que se desean visitar.

En esta sección vamos a presentar nuestra propuesta para modelar las preferencias de los turistas. Para ello descompondremos su perfil en varias estructuras: la principal se basará en el conocimiento del turista formalizado en la ontología propuesta y sus descripciones semánticas; en las otras, incluiremos los puntos valorados por el turista y su

estereotipo⁷⁰. Las técnicas de modelado propuestas deben complementarse con mecanismos de aprendizaje para adaptarse a los posibles cambios de los intereses de los turistas.

La mayoría de las técnicas de modelado de usuarios no se adaptan bien a nuestra propuesta, ya que son excesivamente simples para nuestros objetivos, lo que se traduce en unas capacidades de inferencia insuficientes (en la sección 4.5.1 se han descrito diferentes técnicas de representación de la información del perfil del usuario). El principal inconveniente reside en la elevada complejidad algorítmica de los procesos necesarios para generar resultados de interés, lo que impide que estos procesos se ejecuten en un terminal móvil (obviamente con una potencia de cómputo reducida). Por este motivo, la gestión de los perfiles y de la información de los ítems a recomendar se deben ubicar en la capa del servidor, responsable asimismo de la ejecución de las operaciones de recomendación

El uso de la misma ontología empleada para representar el conocimiento de los puntos turísticos presenta una desventaja, ya que al ser demasiado extensa, la obtención del perfil completo del turista sería un proceso demasiado tedioso para éste. Por este motivo, se propone registrar inicialmente información de las categorías superiores y no de todas las categorías existentes. Posteriormente, a medida que el turista vaya realizando visitas y vaya valorando ítems se irá complementando y actualizando el perfil, que además debe registrar información acerca de cómo afectan los diferentes atributos o características contextuales a las preferencias del usuario.

⁷⁰ El concepto de estereotipo se explicará con detalle en el apartado 5.7.2.

Cuando el usuario se da de alta, el perfil se construirá mediante el uso de técnicas de recopilación de información explícita, creando el denominado *perfil individual*. Esta información se modulará y refinará mediante la interacción con el sistema a través de técnicas de retroalimentación. Para ello se propone utilizar dos métodos de recolección de información:

- a) Explícita: el sistema debe permitir al usuario valorar⁷¹ dentro de un rango determinado los ítems que ha visitado, pudiendo indicar el contexto en el que realiza la valoración.
- b) Implícita: el sistema captura la interacción del usuario de sin intervención directa de éste, como por ejemplo determinando el tiempo que está en un punto, el orden en que visita los ítems recomendados, si acepta o rechaza una recomendación, etc.

Con toda esta información individual y personalizada, el sistema asignará al turista uno de los grupos o estereotipos que el sistema tenga identificados, según la afinidad del turista con dicho grupo. Como ya se ha indicado antes, tanto la información individualizada como la asignación de un usuario a un grupo pueden variar a medida que el usuario vaya utilizando el sistema.

Por ejemplo, un turista podría pertenecer a un estereotipo como el siguiente⁷²:

- DestinoPlaya: 5
- DestinoFamiliar: 5

⁷¹ Como hemos comentado en anteriores ocasiones las escalas pueden ser numéricas o lingüísticas, por lo que se hacen necesaria operaciones de conversión.

⁷² Se incluyen solo algunas características para no hacer tedioso el ejemplo.

- DestinoLagos: 1
- DestinoCiudad: 2
- DestinoMontaña: 1
- DestinoRural: 2
- DestinoRío: 2
- DestinoVerano: 5
- DestinoInvierno: 2
- DestinoCultural: 3

donde se ha utilizando una escala numérica del 1 al 5, donde el 1 significa Nada y el 5 Mucho.

En nuestro modelo se utilizan dos técnicas explícitas para obtener la información relativa al perfil inicial:

- a) El turista valorará las clases superiores de la ontología, profundizando en aquellas categorías en las que muestre mayor interés. Una vez que el turista descarta un tipo, no se pedirá la valoración de subclases inferiores a la rechazada. Por ejemplo, si el turista nos informa de que no le interesa visitar espacios naturales, no habrá que preguntarle por nacimientos de ríos, lagos o miradores. Es evidente que esta forma de solicitar información al turista ahorra tiempo y no sería posible si se utilizaran métodos tradicionales carentes de organización jerárquica.
- b) Para poder categorizar al turista, se utilizan recursos o actividades muy conocidas, con el objeto de identificar sus

preferencias turísticas y características demográficas, asignándole alguno de los estereotipos que el sistema contempla.

De todas formas, el sistema debe permitir que aquellos turistas que lo deseen puedan solicitar que, al inicializar el perfil, puedan valorarse categorías inferiores de la estructura jerárquica.

Cuando se registra el usuario, como hemos dicho anteriormente, se realiza una recolección de datos explícita, la cual es indispensable para que el sistema comience a realizar recomendaciones coherentes respecto a las preferencias del turista. Las técnicas explícitas pueden aplicar diferentes métodos de recolección, entre los que podemos destacar los siguientes ejemplos que consideramos el sistema propuesto debe permitir:

- a) Solicitar al usuario que pondere en base a una escala diferentes tipos de categorías de puntos de interés, como, por ejemplo, la categoría de museos.
- b) Instar al usuario a que ordene un conjunto de temas de una lista. Por ejemplo, se le muestra una lista de tipos de puntos de interés y tiene que ordenarlos de mayor a menor preferencia.
- c) Presentar al usuario un conjunto de opciones para que seleccione solo una, descartando el resto de opciones.
- d) Seleccionar un número de ítems de un listado que se le oferta al usuario.

La recolección de información implícita durante la utilización del sistema complementa la información del turista. Entre los métodos que proponemos que se pueden aplicar destacan los siguientes:

- a) Guardar un registro de las categorías de ítems que el usuario visita, así como del contexto en las que las realiza y las valoraciones que otorga.
- b) Analizar el número de visitas recomendadas por el sistema que han sido realizadas y aquellas que han sido rechazadas.
- c) Guardar un registro del orden de visita.

A continuación se aporta un ejemplo de creación de perfiles individuales, en el que dos turistas se registran y suministran información sobre cinco categorías de puntos turísticos (tabla 5.6) y tres descriptores de turista que nos indican el estereotipo al que está asignado el turista. Se supone por simplicidad en la comprensión del ejemplo, una escala de valoración donde el valor 1 significa nada interesado y el valor 5 indica muy interesado:

Categoría	Turista 1 (u_1)	Turista 2 (u_2)
Edificaciones	3	5
Edificios públicos	2	2
Edificios religiosos	1	5
Restos arqueológicos	5	2
Monumentos	4	5
Descriptor	Turista 1 (u_1)	Turista 2 (u_2)
Turismo deporte	1	5
Turismo naturaleza	2	4
Turismo cultural	4	4

Tabla 5.6. Ejemplo de valoraciones sobre preferencias de dos turistas.

En base a los datos del ejemplo de la tabla 5.6 se pueden aplicar las definiciones 5.22, 5.23 y 5.24, observando que los dos usuarios tienen el mismo interés en visitar edificios públicos, similares en monumentos y dispares en edificios religiosos.

5.7.2. Creación de estereotipos

Cada turista se asignará a un estereotipo el cuál describe las características demográficas (nivel de estudios, nivel cultural, nivel económico, etc.) del turista así como qué tipo de turista es (turista de sol y playa, turista cultural, turista de aventuras, turista de descanso, etc.). La creación y asignación de estereotipos dependerá por tanto del número de instancias que el sistema presente al turista para crear la parte de su perfil grupal, es decir que cada turista estará asignado a uno de los estereotipos que el sistema proporcione. Por ejemplo, si el número de instancias (n) que presenta el destino fueran cuatro (Turismo familiar, Turismo cultural, Turismo de descanso y Turismo de naturaleza), y el número de valoraciones (k) para cada instancia es 3 (Nada, Normal y Bastante), tendremos que el total de estereotipos en este destino se calcula mediante la siguiente expresión:

$$Nº\text{Estereotipos} = k^n = 3^4$$

Para que no sea demasiado compleja la tarea de asignación de estereotipos, la opción más sencilla es presentar al turista una instancia para cada tipo de descriptor turístico y dato demográfico, siendo el número de valores posible pequeño (nosotros proponemos un valor máximo de 5).

De esta forma, cuando se quiera conocer si un determinado punto es interesante para un turista, se pueden utilizar técnicas de recomendación demográficas. Estos métodos proponen analizar las valoraciones que han realizado distintos turistas que pertenecen al mismo estereotipo del turista que solicita la recomendación, o pertenecen parcialmente a otros estereotipos (definiciones 5.25 y 5.26).

El modelo que utilizaremos está basado en las distancias. Si ilustramos el ejemplo de los dos usuarios de la tabla 5.6, cada uno pertenecerá directamente a un estereotipo o grupo de forma completa, pudiendo pertenecer a otros estereotipos de forma parcial. En las siguientes definiciones supondremos que todas las instancias pueden tomar el mismo número de valores, utilizando una escala del 1 al 5⁷³.

Definición 5.25: Dos turistas u_1 y u_2 pertenecen totalmente al mismo estereotipo, si se cumple que tienen la misma valoración para todos los descriptores turísticos y demográficos.

Definición 5.26: Dos turistas u_1 y u_2 pertenecen parcialmente a un estereotipo, si la definición del estereotipo del usuario u_1 respecto al usuario u_2 tiene una diferencia de valoración menor o igual a un valor umbral⁷⁴.

En nuestro caso consideraremos el umbral al número de instancias turísticas-demográficas

⁷³ Formalmente esto no tendría que ser así, pero en esta sección lo hemos considerado de esta manera para facilitar la comprensión. En las definiciones 6.15 a 6.17 se define con más formalidad estos conceptos contextualizándolo a los motores de recomendación.

⁷⁴ El umbral para considerar si pertenece o no parcialmente debe ser configurable.

Así, en el ejemplo de la tabla 5.6 podemos observar que el usuario 1 estará asignado al estereotipo (1, 2, 4) y el usuario 2 formará parte del estereotipo (5, 4, 4). El usuario 1 no pertenece al estereotipo (5, 4, 4), pero sí pertenece parcialmente al estereotipo (3, 2, 4).

5.7.3. Adquisición del perfil de necesidades

Otro mecanismo que utilizaremos en nuestro sistema para poder realizar recomendaciones es recoger las necesidades de recomendación (Balabanovic y Shoham, 1997; Herrera, Herrera-Viedma y Martínez, 2000) actuales del turista por medio de una simple calificación de atributos contextuales y no contextuales.

En muchas situaciones, las necesidades de recomendación pueden variar por muy diversos motivos, por lo que se requiere seleccionar claramente qué tipo de ítem se quiere visitar. Por ejemplo, si tenemos ganas de ir a almorzar, no tendremos ningún interés en que el sistema nos recomiende visitar un museo.

Definición 5.27: Se define necesidad del usuario u y se denota por $N(u)$ a una lista de atributos que el turista empleará para describir sus requisitos actuales de visita. Estos requisitos pueden ser de una categoría o de varias categorías (restaurantes, bares, teatros, etc.)

$$N(u) = (n_1, n_2, \dots, n_k)$$

Este conjunto de atributos relacionados con las necesidades actuales del turista debe coincidir con el utilizado en los ítems a recomendar. Los atributos pueden ser valorados utilizando diferentes escalas lingüísticas

de acuerdo a la naturaleza del propio atributo, mejorando la calidad del conocimiento que tendremos sobre las necesidades del turista.

Esta forma de conocer las necesidades puede resultar tediosa o incómoda para algunos usuarios, por lo que debe permitirse emplear como descripción de sus necesidades actuales seleccionar un ítem recomendado anteriormente. Es decir las características de búsqueda o necesidad actual coincidirá con las características de un ítem visitado con anterioridad, permitiendo al turista un refinamiento de estas necesidades si así lo considera oportuno. De esta forma, se consigue que la adquisición de los requisitos del turista sea más sencilla. Este tipo de técnica tiene como principal ventaja que no requiere utilizar el perfil individual del turista ni el de grupo, permitiéndole expresar sus necesidades en el momento concreto en el que solicita la recomendación.

Un ejemplo de perfil de necesidades de un turista que quiere ir a un restaurante se ilustra en la tabla 5.7:

Característica	Valor lingüístico	Valor numérico
Económico	Bastante	0,75
Local cerrado	Sí	1
Ambiente familiar	Bastante	0,75

Tabla 5.7. Necesidades de un usuario que solicita restaurante

En este ejemplo el turista quiere encontrar restaurantes que sean bastante económicos, que el local sea cerrado y que tenga un ambiente bastante familiar.

5.8. Solicitudes de recomendación

En esta sección vamos a describir las diferentes formas de interacción que la propuesta realizada debería permitir, con el objetivo de conseguir que las recomendaciones obtenidas sean sencillas de obtener por parte del turista. Por otro lado, también realizaremos una breve exposición de los distintos mecanismos que se proponen para recuperar los diferentes puntos de interés que permitan cumplir las expectativas del turista que solicita la recomendación.

5.8.1. Modos de interacción del turista

El objetivo principal del sistema propuesto es la utilización de motores de recomendación contextuales en un sistema de realidad aumentada. Cuando un turista solicita una recomendación se deben proporcionar los mecanismos para que pueda obtenerse el contexto en el que se solicita la recomendación.

Una vez obtenida la información contextual, se debe permitir al turista diferentes formas de interacción con el sistema para poder expresar sus solicitudes de recomendación, que pueden ser agrupadas en:

- a) Interacción libre: el turista no tiene ninguna preferencia especial en ese momento y espera sugerencias del sistema teniendo en cuenta el perfil del turista.
- b) Interacción guiada por categoría: el visitante selecciona una categoría y el sistema debe recomendarle ítems correspondientes a dicha categoría. Además, de seleccionar la categoría en la que está interesado, el turista puede filtrar la recomendación a partir de distintas características que deben poseer los ítems.

c) Interacción de solicitud de itinerario: el turista indica el tiempo que tiene y el sistema genera un itinerario de ítems y tiempos que destinará a esa visita. El origen del itinerario puede ser la posición actual del turista, o una indicada explícitamente por el usuario. Esto último es muy interesante para planificar con antelación un itinerario. Para la recomendación de los ítems que forman el itinerario pueden utilizarse los tipos de interacción indicados en los puntos anteriores..

Por tanto, en todos los tipos de interacción el turista debe tener la posibilidad de expresar también sus necesidades no contextuales a través de una serie de restricciones, para que el sistema pueda realizar sugerencias. Estas necesidades pueden obtenerse directamente del perfil del turista o bien a través de la selección de categorías y características de puntos que desea visitar.

Pongamos un ejemplo real de cada tipo de interacción que suele presentarse muy a menudo cuando un turista se encuentra en una zona de una ciudad, en la que también se especifican atributos contextuales:

¿Qué lugares serían más interesantes visitar esta mañana que está lloviendo? ¿A qué distancia de aquí se encuentran? ¿Estarán abiertos? En estos casos se estaría utilizando la interacción libre, pues se utilizaría el perfil del turista para la realización de las recomendaciones.

Otra variante podría ser, *¿qué restaurantes de comida vegetariana existen en esta ciudad y tienen un ambiente romántico para ir en pareja? ¿Cuáles tienen un precio económico?* En este caso se estaría utilizando la interacción guiada por categoría, dado que se buscan ítems

pertenecientes a la categoría de restaurantes, que tengan entre sus características que ofrezcan comida vegetariana y que tengan un precio económico.

Como podemos comprender, en diferentes momentos las necesidades del turista pueden ser muy diversas. Dado que la mayoría de los sistemas de realidad aumentada muestran diferentes puntos de interés turísticos generales, nuestro objetivo es que pueda recomendar sólo aquellos que le serán más interesantes al turista en el momento de realizar la petición.

5.8.2. Filtrado y recomendación de puntos de interés

A continuación se realiza una exposición de la forma en que el sistema selecciona los puntos de interés que se le recomendarán al turista. El método propuesto combina varias técnicas clásicas de búsqueda. Se trata de un método híbrido (Burke, 2002), en el que pueden intervenir las siguientes fases (no necesariamente en este orden):

- a) Fase de detección de puntos y filtrado contextual: consiste en recuperar todos los puntos de interés cercanos a la posición en la que se encuentra el turista. Según el caso, puede tratarse de un número elevado de puntos de interés, que sería conveniente reducir para optimizar el tiempo de cálculo para las recomendaciones. En este trabajo se propone realizar un proceso de pre-filtrado basado en criterios conceptuales, eliminando aquellos puntos que no verifiquen una serie de criterios. En particular, se propone el uso de técnicas basadas en el análisis

formal de conceptos (AFC-FCA) y que se describe detalladamente en el capítulo siguiente en la sección 6.2.

- b) Fase de filtrado mediante aplicación de restricciones sobre categorías: en esta fase se realiza un filtrado de puntos de interés que cumplan las restricciones de categorías tanto positivas como negativas determinadas previamente por el turista. Si no existe ninguna restricción específica al utilizar el sistema, no se realiza ningún filtrado. Un ejemplo podría ser filtrar todos los ítems que sean museos. En la sección 5.8.2.2 lo describiremos con mayor detalle.
- c) Fase de recomendación basada en el perfil del turista: en esta fase se realiza un estudio de los puntos filtrados, aplicando para ello técnicas que permitirán organizar los distintos puntos de mayor a menor preferencia, según los gustos del turista y las valoraciones existentes. Esta clasificación facilitará al turista enormemente la búsqueda y selección de los puntos que desea visitar. En la sección 6.5 describiremos un modelo de recomendación colaborativo que utiliza el perfil de preferencias del turista.
- d) Fase de recomendación basada en datos demográficos: en esta fase se realiza una predicción del interés que tendrá en el turista los puntos filtrados, teniendo en cuenta las valoraciones realizadas por turistas que han visitado esos puntos y que tienen intereses similares a los suyos. En la sección 6.4 propondremos un modelo de recomendación demográfico que pretende no solo generar recomendaciones basándose en las características

demográficas del turista sino que además introduce un efecto sorpresa para evitar que todos los ítems recomendados sean siempre del mismo tipo.

- e) Fase de recomendación basada en datos históricos: en esta fase se realiza una predicción del interés del turista teniendo en cuenta información extraída de puntos de interés previamente visitados por el propio turista y que fueron bien valorados. En la sección 6.6 se detallará el motor propuesto basado en este tipo de técnicas.
- f) Fase de recomendación basada en necesidades: en esta fase se realiza una predicción del interés partiendo de características que pueden tener los ítems que desea visitar. Es importante destacar que son características de ítems y no de categorías, a diferencia de la fase b. Estas características pueden ser expresadas por el turista explícitamente o extrayendo de forma automática las características un ítem visitado anteriormente y que ha sido seleccionado por el turista porque desea visitar un punto similar. Un ejemplo de petición de recomendación basado en estas técnicas podría ser solicitar sitios que no sean al aire libre, o sitios que tengan aire acondicionado, etc. En la sección 6.3 y 6.7 se describirán motores de recomendación basados en características de los ítems que se desean visitar.

Como se puede observar, se han distinguido dos fases de filtrado y cuatro de recomendación. En primer lugar se deben llevar a cabo las de filtrado y posteriormente las de recomendación, aunque no todas ellas tienen por qué llevarse a cabo. En caso de ejecutarse varias fases de

recomendación se tendrán que aplicar procesos de reordenación mediante técnicas de hibridación (véase sección 3.4).

De hecho, la aplicación de uno u otro modelo de recomendación puede ser configurada por el turista, que podrá solicitar una recomendación indicando un punto de interés visitado anteriormente para que se le recomienden otros similares, aplicando perfil individual, seleccionando puntos de interés bien valorados por otros turistas con características similares, indicando propiedades de lo que desea visitar, etc. Esta capacidad de personalización es bastante enriquecedora, pues el turista no siempre va a recibir recomendaciones aplicando los mismos criterios.

5.8.2.1. Filtrado de puntos utilizando datos contextuales

Normalmente un destino turístico tiene un número bastante grande de puntos turísticos. Obviamente, realizar un proceso de recomendación con un número grande de ítems conlleva que un tiempo de cálculo mayor con un número reducido de ítems. Por tal motivo se hace necesario realizar diferentes etapas de filtrado que permitan disminuir el número de puntos sobre los que se calculará el grado de predicción de interés del turista.

En nuestra propuesta consideramos muy interesante que se definan reglas de implicación que tengan en cuenta atributos contextuales. Por ejemplo si está lloviendo, parece razonable que se visiten puntos turísticos que estén en recintos cubiertos. Por tanto, aquellos puntos que no cumplan las implicaciones se eliminarán y no se tendrán en cuenta para realizar las recomendaciones.

En la sección 6.2 se describe con detalle la forma en la que se realiza este tipo de filtrado. Es importante destacar que estas reglas de implicación deben ser definidas por personal experto del destino turístico.

5.8.2.2. Filtrado de puntos basado en categorías

El filtrado mediante la selección de categorías permite eliminar un importante número de puntos, haciendo que sólo aquellos puntos que cumplan las restricciones definidas por el turista participen en el proceso de cálculo de predicción de interés.

El turista podrá expresar restricciones positivas y negativas, que se corresponden con criterios booleanos. Dichas expresiones pueden definirse utilizando una notación BNF extendida del tipo:

```
<restricción> ::= [NOT] <expresión> { (AND|OR)[NOT] <expresión> }

<expresión> ::= nombre_categoria

<expresión> ::= nombre_categoria <operador> valor_distancia_km

<expresión> ::= (<familia_todas_categorías> <operador> valor_distancia_km)

<operador> ::= (= | > | ≥ | < | ≤ | ≠)

<familia_todas_categorías> ::= "*"
```

Un ejemplo de restricción para una visita podría ser:

(museo) AND (jardines) AND (museo<2)

En este caso el turista estaría expresando que desea que se le recomiendan museos y jardines, estando los museos a menos de dos kilómetros de distancia.

Otro ejemplo de restricción podría ser:

NOT (restaurante) and (*<7)

Esta expresión seleccionaría todos los puntos de interés que no sean restaurantes y que estén a una distancia inferior a 7 kilómetros. Evidentemente el sistema debe proporcionar al turista asistentes que le permitan generar estas peticiones de una forma cómoda e intuitiva.

5.8.2.3. Recomendación de puntos de interés

Una vez que el sistema ha realizado un filtrado de aquellos ítems que no cumplen los atributos contextuales y que satisfacen las restricciones del turista respecto al tipo de categorías o distancias que desea visitar, comienza la fase de recomendación. Anteriormente se ha descrito brevemente que los motores de recomendación que la propuesta utiliza se basan en: perfil del turista individual, visitas anteriores, datos demográficos del turista y/o características de los ítems.

El sistema debe permitir generar recomendaciones utilizando dichos motores de forma independiente o utilizando simultáneamente varios de ellos. En el caso de utilizar varios motores de recomendación será necesaria aplicar técnicas de hibridación genéricas como técnicas de mezclas, pesos o switch (sección 3.4).

En el capítulo 6 se realizará una detallada descripción de los diferentes mecanismos propuestos para generar los resultados que se deben mostrar al turista.

5.9. Casos de uso del sistema

En la sección 5.8.1 hemos descrito las diferentes formas de interacción del turista con el sistema. En esta sección profundizaremos, describiendo las diferentes acciones que el turista podría realizar en el modelo propuesto. Partiremos de que cualquier interacción del turista en el sistema puede ser denominada como una ‘acción’. El sistema propuesto está diseñado de manera que las diferentes acciones que realice el turista retroalimenten el perfil del mismo, mejorando así la precisión de las recomendaciones. A continuación, se enumeran las principales acciones que puede realizar el turista en el sistema:

- a) Recomendación: se trata del proceso principal del sistema propuesto. Se encarga de recoger la petición de recomendación y mostrar una lista de ítems que se ordenará de acuerdo a las preferencias del turista. En estas preferencias se incluyen parámetros tanto de tipo contextual, restricciones positivas, negativas y el perfil del propio turista. El capítulo 6 describe con detalle los diferentes motores de recomendación y filtrado desarrollados.
- b) Selección. Esta acción está relacionada con la aceptación por parte del turista de los ítems recomendados por el sistema. Se trata de una actividad importante pues la aceptación o rechazo de una recomendación influirá en el perfil del turista.
- c) Visibilidad. Esta acción consiste en la visualización de la lista de ítems recomendados en el entorno real mediante el dispositivo de realidad aumentada.

- d) Generación automática de ruta. Consiste en generar una ruta teniendo en cuenta ítems acordes a los requerimientos del turista y que se adaptan a los datos contextuales precisados anteriormente. La obtención de la ruta también tiene en cuenta la trazabilidad de los turistas en el destino. En la sección 6.8 se describe con detalle el mecanismo y las bases teóricas que se utilizan para generar las rutas.
- e) Búsqueda de ítems similares. Consiste en la localización de ítems relacionados con otro seleccionado por el turista y que éste ha visitado anteriormente, permitiendo además añadir o eliminar alguna restricción. Se trata de una funcionalidad muy útil ya que el sistema recomendará al turista ítems con características similares a uno que le ha satisfecho anteriormente. Se trata de una técnica en la cual las características del ítem que se desea visitar no se expresa explícitamente, sino mediante la selección de un punto de interés visitado con anterioridad
- f) Evaluación. Permitirá al turista evaluar un ítem visitado utilizando diferentes criterios⁷⁵, con objeto de que esta información retroalimente el perfil del usuario y la información asociada al ítem visitado. Dicho de otro modo, si un turista utiliza el sistema y no realiza la actividad de evaluación, estará empobreciendo la precisión de las recomendaciones. De ahí que la actividad de evaluación sea primordial para un óptimo funcionamiento del sistema. El sistema propuesto utilizará

⁷⁵ El turista puede realizar una única valoración general de la visita relacionada, o puede realizar unas valoraciones más exhaustivas valorando la visita según diferentes categorías, atributos contextuales y atributos no contextuales.

diferentes tipos de escalas (numéricas y lingüísticas) por lo que se hará necesario el uso de operaciones de conversión de un tipo de escala a otra para poder generar adecuadamente las recomendaciones.

- g) Visita off-line. Es frecuente que un turista tenga planificada la visita a una serie de puntos seleccionados antes de ir al destino, sin haber realizado una petición de recomendación. Por tanto, es necesario que el sistema permita evaluar un ítem que no ha sido recomendado por el sistema.
- h) Configuración de recomendación. Mediante esta acción el turista podrá seleccionar cuáles de los motores de recomendación implementados por el sistema desea que se apliquen para generar las recomendaciones.
- i) Registro de datos contextuales. En esta actividad el turista informa al sistema de aquellos atributos contextuales que desea que se tengan en cuenta al generar la recomendación (Figura 5.7). Algunos de los datos contextuales se capturan de forma implícita, como por ejemplo la temperatura o el día de la semana y otros son modificados de forma explícita como, por ejemplo, la compañía con la que realiza la visita, la distancia máxima en la que puede estar el ítem que va a visitarse, etc.



Figura 5.7. Detalle aplicación indicando características contextuales

- j) Registro de características. El turista puede informar al sistema de las características que deben poseer los puntos que le van a ser recomendados. Se refiere a valores relacionados a características de los diferentes ítems y no a las categorías a las que pertenecen. Por ejemplo, que el ítem a recomendar esté en un recinto cubierto, que disponga de aire acondicionado, etc. Para ello es necesario el uso de una escala (generalmente lingüística) en la que el turista expresará adecuadamente sus necesidades de visita. Dado que el número de características que puede tener un ítem es muy grande, el turista expresará sólo las que considere más importantes.

5.10. Evaluaciones de los turistas

Como ya se ha indicado, uno de los puntos fundamentales del sistema de recomendación propuesto es el proceso de evaluación de ítems que realizan los turistas, ya que permitirá personalizar el sistema y que las recomendaciones se generen de una forma más precisa.

Muchos sistemas de recomendación se basan en las denominadas evaluaciones positivas (Schwab, Kobsa y Koychev, 2001), es decir, para la creación del perfil del turista sólo se tienen en cuenta las evaluaciones que el turista ha realizado de los ítems visitados. En nuestra propuesta se ha enriquecido este método, permitiendo que el perfil pueda actualizarse en diferentes momentos y no sólo en el de la evaluación de ítems.

Se propone que cada ítem tenga su propia evaluación basada en las categorías a las que pertenece y en sus características. Por ejemplo, si un punto turístico está catalogado dentro de tres categorías, la valoración general (definición 5.8) que realizará el turista puede ser realizada de dos formas diferentes:

- a) El turista asigna una única valoración que será asignada a las categorías a las que pertenece dicho ítem.
- b) El turista asigna una valoración diferente para cada una de las categorías a las que pertenece dicho ítem.

Además, aquellos usuarios que lo deseen podrán también calificar la actividad incluso por diferentes características del producto, aunque esto último será opcional (por ejemplo en un restaurante se podrían evaluar características como precio, servicio, calidad de la comida, etc.).

Así, la catedral de Málaga podría no sólo tener una valoración general por parte del turista, sino que podrían ser evaluados por diferentes categorías como edificios religiosos, museo o jardines.

En el sistema propuesto el turista hace una evaluación en dos fases: una primera que consiste en filtrar los resultados generados por los

motores de recomendación, seleccionando los ítems que realmente le interesa visitar; en una segunda fase, el turista evalúa los ítems visitados según categorías y/o características, independientemente de los motores de recomendación utilizados. Un aspecto novedoso del sistema propuesto es que el rechazo o aceptación por parte del turista de un ítem recomendado implica una evaluación en sí.

Más detalladamente, en la primera fase todo ítem recomendado puede ser clasificado en uno de los siguientes niveles:

- El ítem es interesante: el turista selecciona este ítem porque le interesa visitarlo. Al ser aceptado, el sistema exigirá que en caso de ser visitado, debe ser evaluado.
- El ítem ya ha sido visitado: es posible que se recomiende un ítem ya visitado. En este caso el sistema debe permitir la búsqueda de ítems similares si así lo desea el turista, pero excluyendo los ya evaluados.
- El ítem no es relevante: el turista no acepta este punto de interés recomendado porque no se encuentra interesado actualmente en visitar ítems con estas características.
- El ítem no se selecciona: el turista no acepta pero tampoco rechaza el ítem recomendado, por lo que queda en espera por si el turista quiere aceptarlo o rechazarlo en otro momento.

Es evidente que resultaría muy enriquecedor conocer el motivo por el cual el turista clasifica cada punto de interés en un nivel dado, para de esta forma comprender mejor sus gustos. Quizá un mayor número de niveles permitiría obtener una información más rica sobre el motivo por

el cual el turista evalúa el ítem de una u otra forma. Por ejemplo, podría evaluarse de la siguiente manera:

- El ítem es aceptado porque es lo que el turista estaba buscando.
- El ítem es aceptado por el turista aunque lo ha visitado anteriormente, porque el resto no le gusta.
- El ítem no es aceptado porque el turista no tiene tiempo para visitarlo.
- El ítem no es aceptado porque no le gusta al turista.
- El ítem no es aceptado porque alguna característica contextual (distancia, horario, etc.) le impide visitarlo, pero le gustaría poder realizar la visita en otro contexto.
- El ítem no es aceptado por el turista porque aunque se ajusta a sus preferencias, ha escuchado opiniones que le hacen rechazarlo.

Esta propuesta de niveles es mucho más precisa, proporcionando mucha más información sobre el perfil del turista. En teoría emplearla sería una buena opción, pero en la práctica los turistas prefieren formas de evaluación sencillas, por lo que todo aquello que lo haga más complejo con seguridad incidiría en una menor utilización del sistema.

La mayoría de los sistemas que se encuentran en la literatura contemplan la posibilidad de aceptar/rechazar ítems recomendados, pero no almacenan este tipo de información (Fesenmaier et al., 2006). El método propuesto sí registrará en la información asociada tanto al turista como al ítem, si éste fue o no seleccionado, y si posteriormente

fue evaluado, influyendo como se ha indicado en el perfil del turista y en la valoración de los ítems.

Debe tenerse en cuenta que es posible que un mismo ítem (por ejemplo un restaurante) sea evaluado más de una vez por el mismo turista. En este caso, en el sistema propuesto se considera más consistente que la evaluación final sea la media de todas las evaluaciones realizadas por el turista. Otra opción sería almacenar sólo la última calificación, pero no la consideramos apropiada. El motivo de elegir la valoración media es evitar que una experiencia última muy negativa influya y distorsione los diferentes perfiles.

Resulta de especial interés disponer del historial de evaluaciones realizadas por el turista, ya que no solo sirve para retroalimentar su perfil y el del ítem, sino que además le permitirá solicitar ítems similares a otros visitados con anterioridad y que le resultaron altamente satisfactorios.

El uso del historial de visitas por parte del turista, puede tener una amplia aplicación, como por ejemplo:

- Modificar evaluaciones realizadas anteriormente porque se haya producido algún error, o porque se deseé valorar de forma más exhaustiva.
- Buscar puntos de interés turísticos similares a otros que se han visitado y evaluado anteriormente.
- Recuperar búsquedas o peticiones de recomendación realizadas por el turista.

5.11. Funcionamiento general del sistema

En esta sección vamos a realizar una descripción del modo mediante el cual se realizarán las generaciones de recomendaciones, así como los tipos de información que se utilizará en cada uno de los motores de recomendación propuestos. Dado que el modelo es híbrido y podrá utilizar diferentes motores de recomendación simultáneamente, se explicará la forma en la que se generará una única lista de recomendación. Además, se profundizará en aspectos relacionados con la forma en que el turista indica restricciones para la recomendación, y cómo puede afectar al perfil del turista las diferentes evaluaciones, visitas y rechazo de recomendaciones que realice.

5.11.1. Perfil del turista

Como ya hemos descrito en varias ocasiones, cada vez que el usuario se conecta al sistema de realidad aumentada, éste debe detectar la posición del turista y el servidor seleccionará el conjunto de ítems existentes en esa zona. El usuario deberá indicar la forma de interacción (libre, guiada por categoría o solicitud de itinerario⁷⁶) para obtener del sistema de recomendación la lista de sugerencias, según las preferencias del propio turista que están determinadas en su perfil. El subsistema de información contextual realizará un filtrado previo, para reducir el número de puntos sobre los que realizar el cálculo de las predicciones.

Para facilitar la comprensión, a continuación se realizan algunas definiciones en relación a las preferencias del turista que solicita la

⁷⁶ Sección 5.8.1.

recomendación que sirven para complementar otras definidas en secciones anteriores:

Definición 5.28: Las preferencias específicas contextuales del usuario u (denotadas SCP^u) constituyen una descripción de los atributos contextuales que el usuario tiene en un momento determinado. Algunos de estos atributos son obtenidos automáticamente por el sistema y otros, explícitamente indicándolos el propio usuario⁷⁷.

Definición 5.29: Preferencias específicas del usuario u . Desde el punto de vista formal, una petición de recomendación puede estar compuesta por una lista de preferencias específicas del usuario u , denominada SP^u que a su vez se divide en los diferentes tipos de preferencias descritos en las definiciones 5.18, 5.19, 5.20, 5.21 y 5.28.

Por tanto, SP^u representa el perfil completo del turista, que almacena información muy diversa que contiene toda los datos de SRP^u (definición 5.18), SGP^u (definición 5.19), SHP^u (definición 5.20), SPP^u (definición 5.21) y SCP^u (definición 5.28).

En la figura 5.8 se muestra un ejemplo de la diversa información que contiene la estructura SP^u . De nuevo, hemos supuesto para facilitar la comprensión el uso valores numéricos en una escala del 0 al 10.

⁷⁷ Tal como hemos indicado en diferentes ocasiones, los valores pueden estar expresados en diferentes escalas.



Figura 5.8. Preferencias específicas del usuario u

5.11.2. Generación de restricciones

En la definición 5.27 definimos la necesidad de un turista como un conjunto de características que el turista indica explícitamente que deben cumplir las actividades que se le va a recomendar. Estas necesidades podemos diferenciarlas en restricciones positivas y restricciones negativas. A continuación presentamos los siguientes conceptos:

Definición 5.30: Una restricción r para el usuario u es una tupla $(c, v, N(u))$, donde $c \in C$ es una categoría, v la valoración mínima que el turista exige que deben tener los puntos que se le recomiendan, y $N(u)$ es una lista de necesidades (definición 5.27), determinadas por las características deseables del producto por parte del usuario u . Los valores c y v son opcionales, en ese caso se establecería la restricción

formada únicamente por características, independientemente de las categorías de ítems y de las valoraciones que tengan dichos ítems.

El sistema debe gestionar en la lista de necesidades dos tipos de restricciones: las denominadas restricciones positivas (denotadas RP^u), que representan características que deben ser satisfechas por los ítems recomendados y las restricciones negativas (denotadas RN^u), que representan características que no deben ser satisfechas por los ítems recomendados.

En la tabla 5.8 se muestra un ejemplo un turista que impone como restricción un restaurante italiano, ya que le gusta bastante este tipo de comida (8), pero desea que el sitio tenga un ambiente tranquilo, que sea económico y que no sea exterior. Esta última es un ejemplo de una restricción negativa.

Categoría	Grado de satisfacción	Necesidades	
		Característica	Valor
Restaurante Italiano	8	Ambiente tranquilo	Bastante
		Económico	Normal
		Exterior	No

Tabla 5.8. Ejemplo de restricciones de un usuario

En los motores de recomendación que utilicemos restricciones (sección 6.3 y 6.7), por simplicidad los valores c y v , no los usaremos.

5.11.3. Módulos de recomendación

Como hemos comentado anteriormente, RAMCAT incorpora un sistema de recomendación compuesto por varios módulos que serán descritos con detalle en el siguiente capítulo. En esta sección vamos a describir qué tipo de información recopilada del usuario se utilizará para generar las recomendaciones. Los módulos de recomendación incorporados son:

- a) Módulo demográfico, que está basado en la técnica de recomendación demográfica (véase sección 3.2.3), que clasifica a un usuario u en una categoría demográfica o estereotipo, de acuerdo con las preferencias y características generales del usuario (SGP^u). Esta técnica recomienda ítems con buenas valoraciones de aquellos turistas que pertenezcan al mismo estereotipo total (intereses idénticos) o parcialmente (intereses muy similares).
- b) Módulo de recomendación basado en contenido, utiliza la técnica basada en contenido (véase sección 3.2.2), que consiste en recomendar una serie de ítems teniendo en cuenta las características de aquellos valorados positivamente por el turista en el pasado (SHP^u) o bien, las preferencias de planificación específicas (SPP^u), indicando las características de lo que desea visitar, asumiendo que esos items le serán de interés. En el caso de utilización de la estructura SHP^u hablaremos de un sistema de recomendación con memoria, mientras que en el caso de utilizar SPP^u se denomina sistema sin memoria.

- c) Módulo de recomendación colaborativo (véase sección 3.2.1), que utiliza técnicas de filtrado de información de preferencias generales del usuario, que consiste en recomendar ítems que han sido bien valorados por otros usuarios que tienen una alta similitud respecto a las restricciones especificadas en el perfil (SRP^u) del usuario que solicita la recomendación. La principal diferencia con el módulo demográfico es que la similitud entre los usuarios no se mide por las características demográficas de los mismos, sino por las valoraciones dadas.
- d) Módulo de filtrado de preferencias históricas, el cual tiene en cuenta el historial de visitas para realizar las recomendaciones. Son técnicas que trabajan con la estructura (SHP^u). En este módulo no se tienen en cuenta las características de los productos, sino las categorías en las que están asignadas.

Todos estos módulos tienen en cuenta características contextuales, para generar las recomendaciones, de ahí que en todos ellos se emplee la estructura SCP^u. En la figura 5.9 podemos observar cómo se obtienen las diferentes recomendaciones en el sistema, partiendo de las diversas fuentes de información incluidas en el perfil del turista:

5.11.4. Módulo de comparación

Los motores de recomendación reciben como entrada de datos una lista de ítems próximos a la posición del usuario y el perfil del usuario que solicita la recomendación. Previamente esos ítems serán filtrados por restricciones de categorías y por el sistema de pre-filtrado contextual (sección 6.2), eliminando aquellos ítems que no cumplen las

implicaciones determinadas por el contexto del usuario. De esta forma, se conseguirá reducir drásticamente el número de ítems que forman el conjunto de entrada a los motores de recomendación. Así, por ejemplo si está lloviendo y queremos ir a un restaurante, evidentemente aquellos que sean exteriores no son una buena opción.

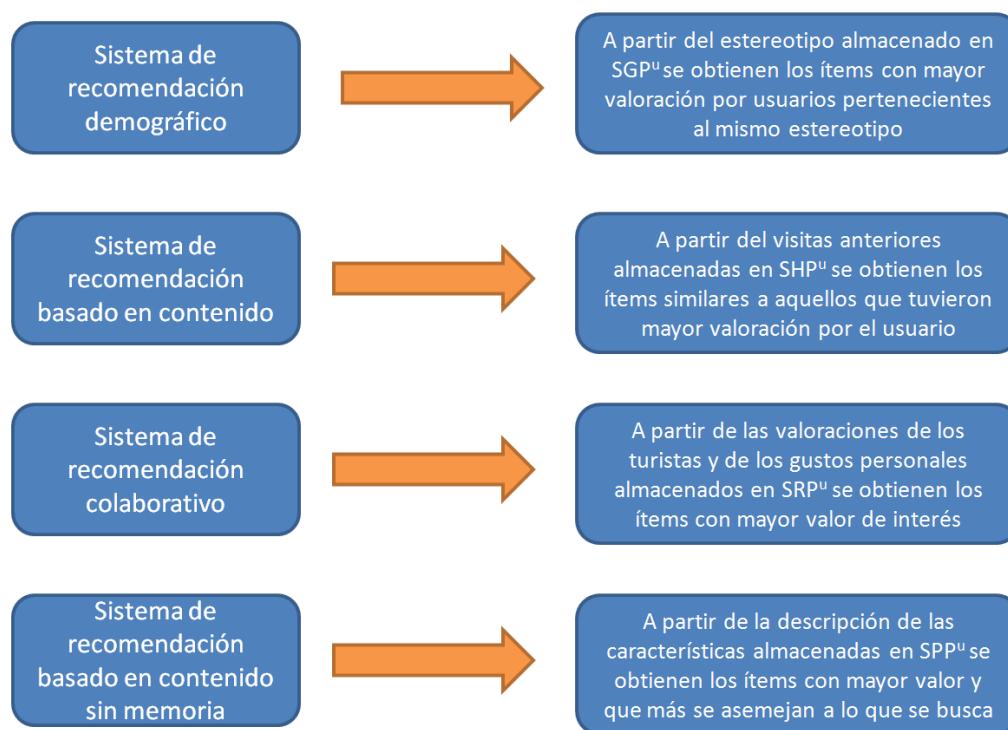


Figura 5.9. Motores de recomendación de la propuesta

Una vez que cada motor genera sus recomendaciones, estos datos deben ser la entrada al módulo de comparación integrado en el subsistema de gestión de recomendación de puntos, responsable de generar la lista definitiva de ítems a recomendar.

Un ítem i aparecerá en la lista de recomendaciones si está en la lista generada por algunos de los motores y cumple las siguientes propiedades:

- a) No pertenece a alguna de las categorías que el turista no desea visitar.
- b) No ha sido visitado (evaluado) con anterioridad por el turista.

Esta propiedad puede ser deshabilitada por el usuario, ya que en algunos casos (para la recomendación de restaurantes, por ejemplo) puede ser admisible repetir la visita.

Llegado a este punto, el sistema ya tiene disponible una lista conjunta de ítems a recomendar. A continuación se describe cómo el módulo de comparación puede combinar los resultados de cada uno de los motores utilizando para ello diferentes técnicas de hibridación (véase la sección 3.4). De las diferentes técnicas existentes, el sistema ofrece al turista la posibilidad de seleccionar una de estas dos:

- a) Cascada: la entrada de un motor de recomendación es la salida del otro. Por tanto en este caso no se realiza una comparación propiamente dicha, sino que los motores se ejecutan en cascada, es decir, cada motor actúa como un filtro del conjunto de ítems generando un conjunto de entrada reducido para el siguiente motor en ejecutarse.
- b) Mediante pesos: se acumulan las valoraciones obtenidas en cada uno de los motores de recomendación y se ordenan los ítems de la lista conjunta en orden decreciente según las valoraciones acumuladas.

5.11.5. Retroalimentación de perfiles

Cuando un turista acepta una recomendación, deberá evaluar posteriormente el ítem visitado. Como ya se ha indicado, la valoración

influirá tanto en la información asociada a los ítems que visita como en su perfil individual y en el de su estereotipo.

Es importante destacar que las escalas que se utilizan para valorar una visita pueden ser muy diversas, pero en el caso de utilizar escalas lingüísticas es indispensable el uso de operaciones de transformación a un número entero⁷⁸.

La retroalimentación que proponemos consiste en que si un turista valora un ítem que pertenece a una categoría con un valor mayor que el que tiene asignado en esa categoría en su perfil, éste debería ir cambiando de forma que aumente la preferencia que tiene el turista en ese tipo de categoría. De forma similar sucedería si valoramos negativamente una visita.

Por ejemplo, supongamos una escala de valoración comprendida entre el 1 y el 5 y tenemos un turista que en su perfil indica que las catedrales le gustan con un valor 3. Si al realizar una visita a una catedral, valora la visita con un 5, el valor del turista en la categoría catedral tendría que ir aumentando.

En un caso genérico, supondremos que n es el número de diferentes valores de la escala que utiliza el turista para realizar una valoración y r es el número de visitas que pueden hacer cambiar la valoración que un turista tenga en una determinada categoría, es decir, este valor r será un valor configurable que cuantifica el número de valoraciones de ítems

⁷⁸ En el capítulo 6 se describe y formaliza con detalle la transformación de una escala lingüística a una escala numérica entera entre 1 y el número de valores lingüísticos de la escala.

necesario para aumentar o disminuir un nivel en la escala de la valoración que tiene un turista en una categoría o un descriptor.

Supondremos que m es el valor central de la escala (asumimos un número impar de valores en la escala). En el caso de que el número de valores de una escala fuera par el mecanismo de cálculo sería análogo pero haciendo referencia a los dos valores centrales.

La forma de obtener el nuevo valor V_u^c para cada categoría a la que pertenezca el ítem visitado por el turista u sería:

$$\text{Si } v_u^i > m \Rightarrow V_u^c = \min\left(n, V_u^c + \frac{1}{|n - \text{calificación} + 1| \cdot r}\right)$$

$$\text{Si } v_u^i < m \Rightarrow V_u^c = \max\left(1, V_u^c - \frac{1}{\text{calificación} \cdot r}\right)$$

$$\text{Si } v_u^i = m \Rightarrow V_u^c = V_u^c$$

En caso de que el turista no acepte visitar el ítem, entonces:

$$V_u^c = V_u^c - \frac{1}{(m - 1) \cdot r}$$

A continuación, para facilitar la comprensión de las expresiones descritas vamos a poner un ejemplo basado en una escala numérica entera de uno a cinco, donde el usuario para calificar una actividad distingue solamente los siguientes valores: 1, 2, 3, 4 y 5. El hecho de hacerlo positivamente (4 o 5) o negativamente (1 o 2) tendrá influencia en el perfil del turista. El rechazo explícito a la visita de un ítem recomendado, supondremos por defecto equivalente a una valoración

negativa⁷⁹ (2), aunque esta valoración no debe afectar al ítem sino solo al perfil, ya que en caso contrario podría desvirtuar la valoración que los turistas tienen de un determinado ítem.

En el ejemplo vamos a suponer que los valores n y r equivalen a cinco, es decir que cinco calificaciones muy positivas harán cambiar la valoración del turista en una categoría variando su valor actual k^{80} al valor siguiente en la escala.

A continuación, vamos a describir qué incidencias tienen las valoraciones en el perfil individual en la escala de cinco valores:

- a) Supongamos que a un turista le ha sido recomendado una actividad que pertenece a la categoría c y la visita le ha resultado muy satisfactoria. En caso de que haya valorado el ítem con un 5, automáticamente la valoración de la categoría c por parte del turista pasa a ser $k+1/r$ aunque para realizar las recomendaciones se continuará usando el valor k^{81} . Cuando el incremento alcance el valor $k+(n-1)/n$, se le cambiará su valoración en esa categoría por el siguiente, es decir $k+1$, siempre que k no sea el valor máximo, ya que obviamente no se podría aumentar su valor.

⁷⁹ Hemos tomado la decisión de no dar la valoración más baja posible (el valor 1) debido a que el rechazo de una visita puede estar motivada por muchas razones, no solo que no le interese o le guste, por lo que hemos considerado utilizar un valor menos drástico y más conservador, en este caso el valor 2, pero este aspecto puede ser configurable.

⁸⁰ El parámetro k será el valor que tendrá definido en el perfil del turista una determinada categoría (V_u^c)

⁸¹ No olvidar que aunque internamente se cambie el valor de k , para los cálculos que se hacen en los motores de recomendación se utilizarán los valores lingüísticos correspondientes.

- b) Supongamos que a un turista le ha sido recomendado una actividad que pertenece a la categoría c y la visita le ha resultado positiva, valorando la actividad con un 4. En este caso, automáticamente la valoración de la categoría c por parte del turista pasa a ser $k+1/(2^r)$, aunque para realizar las recomendaciones se continuará usando el valor k. Cuando el incremento supere los $k+(n-1)/n$, se le cambiará su valoración en esa categoría por $k+1$.
- c) Supongamos que a un turista le ha sido recomendado una actividad que pertenece a la categoría c y la visita le ha resultado muy negativa (valoración 1). En este caso automáticamente la valoración de la categoría c por parte del turista pasa a ser $k - 1/r$, aunque para realizar las recomendaciones se continuará usando el valor k. Cuando el decremento sea inferior a $k-(n-1)/n$, se le cambiará su valoración en esa categoría por la anterior, es decir, $k-1$, siempre que k no sea el valor mínimo, ya que obviamente no puede tomar un valor inferior.
- d) Supongamos que a un turista le ha sido recomendado una actividad que pertenece a la categoría c y la visita le ha resultado negativa (valoración 2). En este caso automáticamente la valoración de la categoría c por parte del turista pasa a ser $k - 1/(2^r)$, aunque para realizar las recomendaciones se continuará usando el valor k. Cuando el decremento sea inferior a $k-(n-1)/n$, se le cambiará su valoración en esa categoría por un $k-1$.

Por tanto para el ejemplo tendríamos que⁸²:

$$\forall c \in C \text{ donde } c \text{ incluida categorias } (i), V_u^c = \begin{cases} V_u^c + 0.2 & \text{si } v_u^i = 5 \\ V_u^c + 0.1 & \text{si } v_u^i = 4 \\ V_u^c & \text{si } v_u^i = 3 \\ V_u^c - 0.1 & \text{si } v_u^i = 2 \\ V_u^c - 0.2 & \text{si } v_u^i = 1 \end{cases}$$

Si por el contrario una recomendación no es aceptada, la influencia sería:

$$\forall c \in C \text{ donde } c \text{ incluida categorias } (i), V_u^c = V_u^c - 0.1$$

Este proceso se llevará a cabo en todas las categorías a las que pertenezca el ítem que ha visitado el turista.

Un caso especial sucede al visitar y valorar una actividad que pertenece a una categoría que no ha sido valorada en el perfil del turista previamente. En este caso, automáticamente se le asignará el valor de esa primera valoración, que posteriormente irá modificándose de la forma descrita anteriormente.

A continuación se expone cómo una valoración influye en un estereotipo de un turista (sólo variarían los descriptores⁸³ turísticos, no los demográficos). De forma general para cualquier escala, tendríamos que para todos los descriptores turísticos:

$$\text{Si } v_u^i > m \text{ y } d_k^u < d_k^i \Rightarrow d_k^u = \min\left(n, d_k^u + \frac{1}{|n - v_u^i + 1| \cdot r}\right)$$

⁸² Por simplificación suponemos que se hace una valoración general. En caso de que el turista valore el ítem para cada una de las categorías, habría que recalcular el valor V_u^c de forma independiente para cada una de las categorías.

⁸³ El concepto descriptor fue presentado en la definición 5.15.

$$\text{Si } v_u^i > m \text{ y } d_k^u \geq d_k^i \Rightarrow d_k^u = \max\left(1, d_k^u - \frac{1}{|n - v_u^i + 1| \cdot r}\right)$$

$$\text{Si } v_u^i < m \text{ y } d_k^u = d_k^i \Rightarrow d_k^u = \max(1, d_k^u - 1/(v_u^i \cdot r))$$

en el resto de los casos $d_k^u = d_k^i$

Si una recomendación es explícitamente rechazada, la influencia sería:

$$\forall d_k \in D, d_k^u = d_k^u - \frac{1}{(m-1) \cdot r} \text{ si } d_k^u = d_k^i$$

Si se consulta información sobre una determinada actividad que no ha sido recomendada:

$$\forall d_k \in D, d_k^u = d_k^u + \frac{1}{(m-1) \cdot r} \text{ si } d_k^u < d_k^i$$

$$\forall d_k \in D, d_k^u = d_k^u - \frac{1}{(m-1) \cdot r} \text{ si } d_k^u > d_k^i$$

Para la mejor comprensión de las expresiones, pongamos el ejemplo basado en los valores de n y r igual a 5, teniendo en cuenta que nunca se puede superar el valor máximo ni tomar un valor inferior al valor mínimo:

- a) Si un turista valora muy positivamente o positivamente (valor cinco o cuatro en el ejemplo) una actividad que le ha sido recomendada, los descriptores turísticos de la actividad que tengan un valor idéntico al del turista no se modificarán. Aquellos que no coincidan comenzarán a oscilar $k \pm 1/r$ o $k \pm 1/(2*r)$, dependiendo de si la valoración ha sido 5 u 4 y si el valor del descriptor del punto turístico es mayor o menor al del

turista. Para los cálculos se utilizará el mismo factor k hasta que se alcance el valor $k \pm (n-1)/n$. En ese caso, el parámetro del descriptor pasará a $k \pm 1$, cambiando el turista de estereotipo.

- b) Si un turista valora muy negativamente (1) un ítem que le ha sido recomendado por su estereotipo, aquellos descriptores turísticos asociados al estereotipo del turista que tengan el mismo valor del ítem visitado, deberán modificarse a un valor $k - 0,2$, aunque seguirá teniéndose en cuenta el valor k para los cálculos hasta que disminuya al menos $k - (n-1)/n$. Una vez que se supere ese límite, el valor de ese descriptor pasará a $k - 1$, cambiando al turista de estereotipo.
- c) Si un turista valora negativamente (2) una actividad que le ha sido recomendada por su estereotipo, aquellos descriptores turísticos asociados al ítem visitado que tengan el mismo valor del estereotipo del turista, deberán modificarse a un valor $k - 1/(2^*r)$, aunque seguirá teniéndose en cuenta el valor k para los cálculos hasta que disminuya al menos $k - (n-1)/n$. Una vez que se alcance ese límite, el valor de ese descriptor pasará a valer $k - 1$, cambiando al turista de estereotipo.

Por tanto, siguiendo la escala del ejemplo, si una visita es calificada, los descriptores tomarían los siguientes valores:

$$\forall d_k \in D, d_k^u = \begin{cases} d_k^u + 0.2 & v_u^i = 5 \text{ y } d_k^u < d_k^i \\ d_k^u + 0.1 & v_u^i = 4 \text{ y } d_k^u < d_k^i \\ d_k^u - 0.2 & v_u^i = 5 \text{ y } d_k^u > d_k^i \\ d_k^u - 0.1 & v_u^i = 4 \text{ y } d_k^u > d_k^i \\ d_k^u - 0.1 & v_u^i = 2 \text{ y } d_k^u = d_k^i \\ d_k^u - 0.2 & v_u^i = 1 \text{ y } d_k^u = d_k^i \\ d_k^u & \text{en el resto de los casos} \end{cases}$$

Si una recomendación es explícitamente rechazada, la influencia sería:

$$\forall d_k \in D, d_k^u = d_k^u - 0.1 \text{ si } d_k^u = d_k^i$$

Si se consulta información sobre una determinada actividad que no ha sido recomendada:

$$\forall d_k \in D, d_k^u = d_k^u + 0.1 \text{ si } d_k^u < d_k^i$$

$$\forall d_k \in D, d_k^u = d_k^u - 0.1 \text{ si } d_k^u > d_k^i$$

A continuación, se aporta un ejemplo en el que se supone un grupo de tres turistas que visitan durante cinco días un destino turístico. Se tomarán como referencia $n=5$ y $r=5$. Supongamos que se le recomiendan siete ítems cada uno de los días, de los cuáles visitarán las cinco primeras. Supongamos los siguientes valores de los descriptores turísticos (d_1 , d_2 y d_3) de los turistas (Tabla 5.9) y de los ítems recomendados (Tabla 5.10):

Turista	d1	d2	d3
u1	3	4	4
u2	5	4	4
u3	1	5	5

Tabla 5.9. Valor de descriptores turísticos de un grupo de turistas.

	d1	d2	d3																
Día 1				Día 2				Día 3				Día 4				Día 5			
i1	1	3	3	i1	5	4	2	i1	5	3	5	i1	2	5	1	i1	2	5	1
i2	5	1	1	i2	3	3	5	i2	1	1	4	i2	5	1	4	i2	5	5	5
i3	4	4	4	i3	5	2	5	i3	1	3	3	i3	2	4	2	i3	3	1	3
i4	5	5	5	i4	5	5	1	i4	3	3	2	i4	3	4	1	i4	5	5	2
i5	4	5	1	i5	4	2	4	i5	1	2	2	i5	5	4	3	i5	2	1	3
i6	3	2	3	i6	3	4	4	i6	4	4	4	i6	1	4	3	i6	3	3	5
i7	4	5	4	i7	2	4	3	i7	4	2	4	i7	4	4	4	i7	2	3	1

Tabla 5.10. Valor de descriptores turísticos de actividades

recomendadas.

	u1	u2	u3																
Día 1				Día 2				Día 3				Día 4				Día 5			
i1	4	2	1	i1	5	4	4	i1	3	5	4	i1	2	4	5	i1	4	3	4
i2	3	4	2	i2	5	4	5	i2	3	4	3	i2	2	2	4	i2	3	1	2
i3	4	5	1	i3	5	3	3	i3	4	4	2	i3	2	5	2	i3	5	5	2
i4	4	1	2	i4	4	3	2	i4	4	1	4	i4	3	2	1	i4	5	4	5
i5	1	3	2	i5	5	1	3	i5	1	3	3	i5	2	3	3	i5	1	3	3

Tabla 5.11. Valoración de cada turista en cada actividad visitada.

Partiendo de las valoraciones realizadas por los diferentes turistas a las visitas realizadas (tabla 5.11), se aporta una comparación de la evolución en el valor de los descriptores turísticos de los diferentes turistas al comienzo de cada día y al finalizar el viaje (tabla 5.12). En el ejemplo se compara también las diferencias que se producen considerando una retroalimentación compuesta por una sola fase (teniendo en cuenta solo las valoraciones de los ítems visitados) o por dos fases (teniendo en cuenta los rechazos de visitas y las valoraciones realizadas sobre los ítems visitados):

	d1	d2	d3		d1	d2	d3
Sólo evaluaciones				Rechazos + Evaluaciones			
Antes de comenzar el viaje							
u1	3	4	4	u1	3	4	4
u2	5	4	4	u2	5	4	4
u3	1	1	5	u3	1	1	5
Después del primer día							
u1	3	4	4	u1	3	4	4
u2	5	4	4	u2	5	4	4
u3	1	1	5	u3	1	1	5
Después del segundo día							
u1	3	4	4	u1	3	<u>3</u>	4
u2	5	4	4	u2	5	4	4
u3	1	1	5	u3	1	1	5
Después del tercer día							
u1	3	4	4	u1	3	3	4
u2	5	4	4	u2	5	<u>3</u>	<u>3</u>
u3	1	1	5	u3	1	1	5
Después del cuarto día							
u1	3	<u>3</u>	4	u1	3	3	<u>3</u>
u2	<u>4</u>	4	4	u2	<u>4</u>	3	3
u3	<u>2</u>	<u>2</u>	<u>4</u>	u3	<u>2</u>	<u>2</u>	<u>4</u>
Al finalizar el viaje							
u1	<u>4</u>	3	<u>3</u>	u1	3	3	3
u2	4	4	<u>3</u>	u2	4	3	3
u3	2	2	4	u3	2	2	4

Tabla 5.12. Comparación retroalimentación del perfil de los descriptores turísticos⁸⁴.

⁸⁴ Aparece en subrayado el cambio de valor de descriptor

A continuación vemos en las figuras 5.10, 5.11 y 5.12 la evolución que tiene en el usuario 1 la retroalimentación de su perfil en los tres descriptores teniendo en cuenta los dos métodos de retroalimentación:

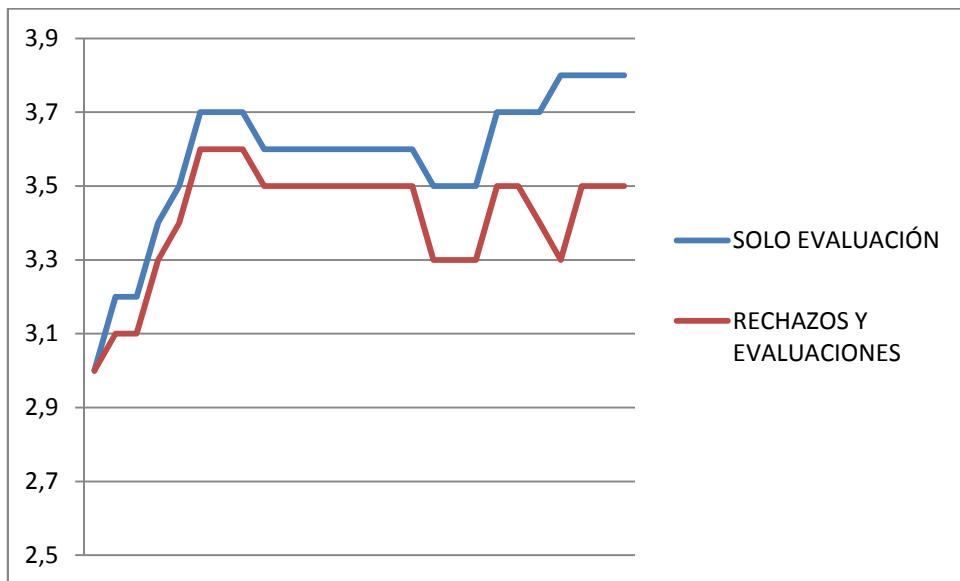


Figura 5.10 Evolución del descriptor turístico d1 en el usuario u1

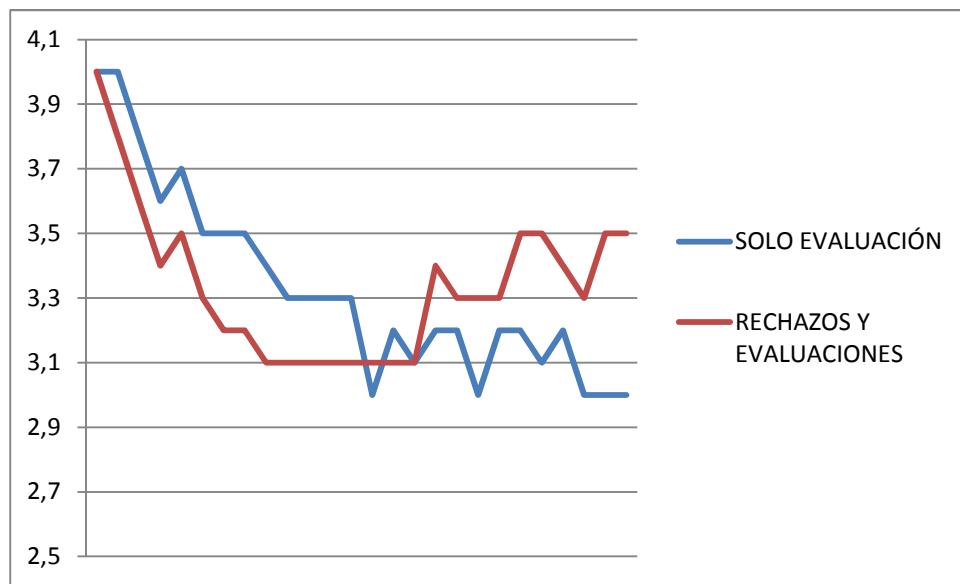


Figura 5.11 Evolución del descriptor turístico d2 en el usuario u1

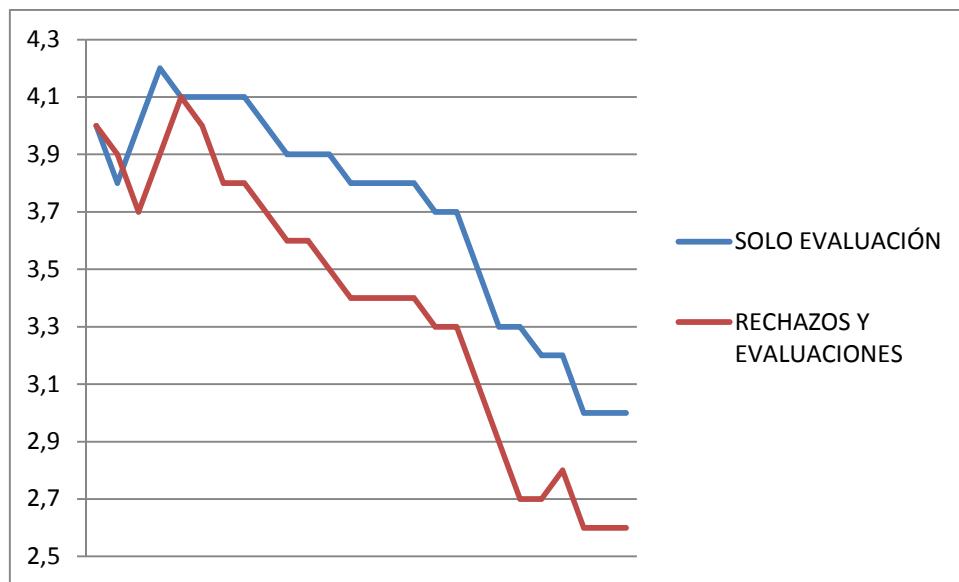


Figura 5.12 Evolución del descriptor turístico d3 en el usuario u1

Aunque la evolución es muy similar, el hecho de contar también con los rechazos como una fuente de información para retroalimentar el perfil, implica que los cambios en los descriptores se realizan de forma más rápida que si no se tiene en cuenta esta información.

Como describimos en la sección 5.10, el hecho de retroalimentar el perfil teniendo en cuenta la calificación que el turista asigne al ítem y también cuando no acepte explícitamente una recomendación, creemos que hará que el perfil esté más ajustado a las preferencias reales. El hecho de rechazar una actividad supone que ese tipo de actividad no se quiere visitar, por lo que esa información debe quedar “registrada” en las características del perfil. De forma similar se retroalimentarán las diferentes características del perfil del turista.

Capítulo 6

Modelos de recomendación

6.1. Introducción

Las diferentes técnicas de recomendación ofrecen resultados correctos, pero nuestra propuesta es ambiciosa y pretende que los resultados obtenidos por las recomendaciones sean los que realmente demanda el turista, por lo que consideramos muy importante tener en cuenta su contexto.

Para ello, tal como hemos descrito en el capítulo anterior se propone como solución el uso combinado de diferentes motores de recomendación, así como un sistema de pre-filtrado contextual que permitirá reducir el número de ítems que forman el conjunto de entradas de los diferentes motores de recomendación. El objetivo final es asegurar que los elementos recomendados se adecuen lo mejor posible al contexto, gustos y preferencias del turista.

Para alcanzar este objetivo se propone el uso de un modelo híbrido contextual, que combina diferentes tipos de técnicas ya descritas en el capítulo 3. El modelo pretende ser flexible, por lo que la activación o desactivación de alguno de los motores dependerá de las características de las recomendaciones que el turista pretenda obtener.

Para poder aplicar el modelo propuesto será necesario disponer de información muy diversa de cada ítem (valoraciones, características, tipos de turistas que visitan un ítem, etc.), lo que posibilitará el empleo de los diferentes tipos de técnicas

En la mayoría de las técnicas de recomendación existentes (Adomavicus y Tuzhilin, 2005; Fesenmaier et al., 2006; Jakkilinkki et al., 2007; Wang y Kao, 2013), los atributos que describen los objetos a recomendar suelen ser los mismos para todos los ítems, mientras que en nuestra propuesta, como hemos descrito en el capítulo anterior, tenemos atributos que dependerán de la categoría a la que pertenezcan el ítem a recomendar, lo que supone una diferencia destacable. Así, por ejemplo, la característica calidad de servicio puede asociarse a un restaurante, pero no a la visita de una catedral.

En el presente capítulo se describen detalladamente los diferentes motores de filtrado, recomendación y generación de rutas que el sistema propone. Se aporta una descripción detallada de cada módulo, formalizando diferentes aspectos relacionados con el mismo, así como un ejemplo que permita comprender mejor su funcionamiento y el tipo de resultados que se genera. Los módulos o motores que se describen son los siguientes:

- Módulo de pre-filtrado contextual utilizando análisis formal de conceptos difuso (PFCBI)
- Motor de recomendación basado en contenido sin memoria utilizando la estructura SPP^u.
- Motor de recomendación basado en técnicas demográficas, empleando la estructura SGP^u.
- Motor de recomendación basado en técnicas colaborativas aplicando la información contenida en la estructura SRP^u.
- Motor de recomendación híbrido basado en historial donde se utilizará la estructura SHP^u.
- Motor de recomendación grupal. Se aplica el motor de recomendación basado en contenido sin memoria para mostrar el funcionamiento de las recomendaciones para grupos.
- Módulo de creación de rutas basándose en la trazabilidad de los turistas y actividades recomendadas al mismo que son seleccionadas para visitar.

6.2. Pre-filtrado contextual utilizando análisis formal de conceptos difuso (PFCBI)

Uno de los principales problemas que tienen los sistemas de recomendación en un destino turístico es el gran número de puntos que pueden visitarse. Cualquier destino consolidado puede tener miles de puntos, por lo que la cantidad de operaciones a realizar para obtener una recomendación es muy elevada. Se hace necesario disponer de

mecanismos que permitan realizar un filtrado que reduzca significativamente el número de puntos que serán entrada a los motores de recomendación (Figura 6.1).

Debido a que en este trabajo nos hemos propuesto incorporar información contextual de manera uniforme, creemos que sería una buena solución utilizar esta información para conseguir ese pre-filtrado. El modelo que vamos a describir está basado en lógica difusa y análisis formal de conceptos (sección 3.6) como marco de trabajo para combinar la información contextual con los puntos turísticos a recomendar.



Figura 6.1. Reducción del número de ítems mediante el proceso PFCBI

6.2.1. Descripción

La gran cantidad de ítems que pueden ser recomendados es un problema para la mayoría de los sistemas de recomendación, pues son muchos los cálculos que deben realizarse en estas situaciones. Otro de los inconvenientes de los sistemas de recomendación en los destinos turísticos es que no suelen utilizar atributos contextuales como el clima, compañía, etc. (Fesenmaier et al., 2006; Adomavicius et al., 2010).

Para mejorar las recomendaciones, debe tenerse en cuenta la información contextual y utilizarla como una categoría adicional de datos. En (Adomavicius y Tuzhilin, 2011) se afirma que los sistemas de recomendación deben tener en consideración tres dimensiones (usuarios, ítems y contexto) y se proponen, como hemos descrito en el capítulo tercero, tres tipos de paradigmas de sistemas de recomendación contextuales:

- a) Pre-filtrado contextual.
- b) Post-filtrado contextual.
- c) Modelado contextual.

Como se ha indicado, en nuestra opinión, un sistema de recomendación para un destino turístico consolidado (probablemente con miles de puntos de interés) debe aplicar el paradigma de pre-filtrado contextual. De esta forma, el sistema de recomendación trabajará con un número menor de ítems, aumentando la eficiencia a la hora de generar los resultados. Otra importante ventaja de esta aproximación reside en la posibilidad de combinarla con cualquiera de las técnicas de

recomendación en dos dimensiones⁸⁵ (usuarios e ítems) existentes. Al tipo de pre-filtrado que hemos desarrollado lo denominamos Pre-filtrado Contextual Basado en Implicaciones (PFCBI).

Muchos autores (Zenebe y Norcio, 2009) proponen el uso de la lógica difusa como base formal para los sistemas de recomendación. Sin embargo, la propuesta que presentamos es novedosa porque nos permitirá cubrir cuestiones propuestas en (Adomavicius y Tuzhilin, 2005) como la incorporación de diversos tipos de información contextual en los procesos de recomendación. Nuestra aproximación se basa en el uso del análisis formal de conceptos (AFC).

Desde el punto de vista filosófico, un concepto es una idea general que corresponde a alguna clase de entidad y que puede caracterizarse por algunas de las características esenciales de la clase. Cuando B. Ganter y R. Wille (Wille, 1982; Ganter y Wille, 1999) conciben un marco dentro de la teoría reticular para la formalización de conceptos, es probable que no adivinaran la amplia difusión de su trabajo original.

Hoy en día, el análisis formal de conceptos (AFC) se ha convertido en un marco de trabajo útil tanto en lo teórico como en las áreas aplicadas. Los trabajos relacionados con AFC cubren áreas como análisis de datos, recuperación de información, representación del conocimiento, etc., considerándose una herramienta muy útil en entornos emergentes como la minería de datos, web semántica, etc.

El principal objetivo del análisis formal de conceptos (AFC) es identificar en una tabla binaria las relaciones entre un conjunto de

⁸⁵ Concepto descrito en la sección 3.3.2

objetos y un conjunto de atributos (se muestra un ejemplo en la tabla 6.1). Estas relaciones establecen una conexión de Galois⁸⁶, que nos permite identificar los conceptos utilizando un marco formal dentro de la teoría reticular. Aparte de la construcción del concepto de retículo, uno de los principales problemas de esta técnica es extraer el conjunto de implicaciones de los atributos en la red reticular. La creación de implicaciones proporciona una información importante, que es extraída en una etapa separada a partir de los datos, formando una representación dual del propio retículo. Una de las ventajas más importantes en el uso de implicaciones radica en la posibilidad de ser manejadas utilizando lógicas de dependencias funcionales (Armstrong, 1974).

	Diseño	
	Espacio abierto	Espacio cerrado
Restaurante estándar	0.3	0.8
Estrellas Michelín	0.1	0.8
Burger	0.3	0.8
Tapas Bar	0.3	0.8
Pizzeria	0.1	0.9
Pescado fresco en playa	0.9	0.2

Tabla 6.1. Representación en AFC de unos restaurantes y su diseño

Una novedad que se presenta en este trabajo es la integración del contexto⁸⁷ en un método basado en AFC por medio de un conjunto de

⁸⁶ Una conexión de Galois es una correspondencia particular entre dos conjuntos parcialmente ordenados.

⁸⁷ En la sección 3.3.1 se define el contexto.

implicaciones. Para ello, proponemos la generación de un conjunto de implicaciones difusas que se corresponden con un contexto dado. Así, cuando el usuario identifica su contexto (compañía, clima, etc.), el sistema enriquece la especificación mediante la adición de un nuevo conjunto nuevo de implicaciones correspondientes a ese contexto. Posteriormente, esa nueva información es tratada por medio de una lógica difusa, que automáticamente reducirá las especificaciones mediante la eliminación de redundancias (Belohlavek et al., 2012). La reducción en un conjunto de implicaciones permite un proceso de validación más eficiente que permite filtrar el conjunto original de ítems, para posteriormente aplicar sistema de recomendación 2D que trabajará con un menor número de ítems. En la siguiente figura se muestra una abstracción del proceso propuesto:

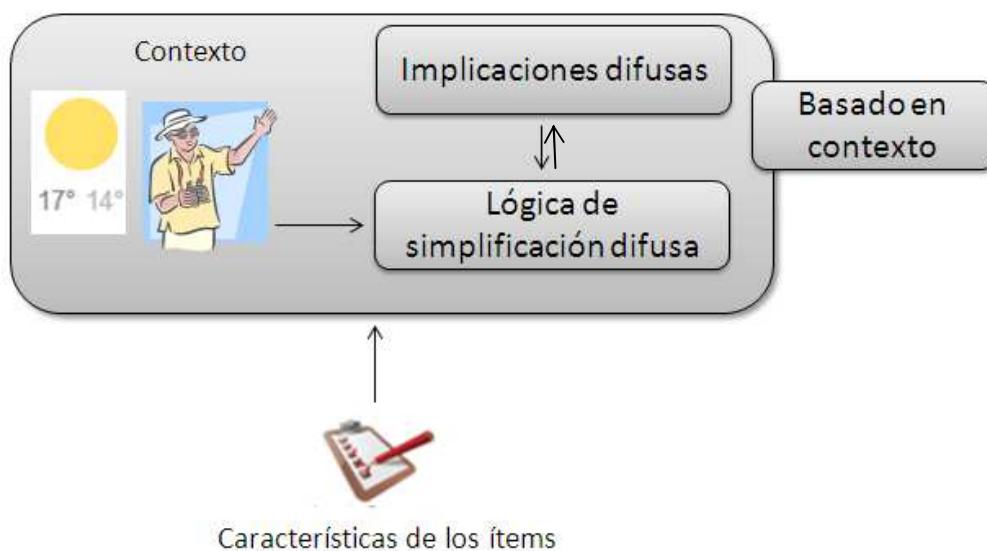


Figura 6.2. SR contextual basado en implicaciones difusas

Algunos autores (Zenebe y Norcio, 2009) han propuesto el uso de la lógica difusa como un marco de trabajo adecuado para la generación de

recomendaciones que soluciona diferentes problemas descritos en (Adomavicius y Tuzhilin, 2005). Ellos proponen un método que permite incorporar información subjetiva, imprecisa o difusa, aspectos que habitualmente aparecen entre las características de los ítems y en la información que suministran los usuarios. Un resultado destacado de este trabajo consiste en que, a pesar de utilizar un lenguaje flexible y Enriquecido para especificar el interés de los usuarios y las características de los ítems, desarrolla un método para inferir recomendaciones que mejora la precisión de éstas.

Algunos autores han utilizado métodos basados en AFC como una aproximación interesante en diferentes sistemas de recomendación. En (du Boucher-Ryan y Bridge, 2006), los autores proponen el AFC como una técnica que permite agrupar ítems y usuarios en conceptos. Este trabajo puede ser considerado como un sistema de recomendación colaborativo y muestra como la aplicación de técnicas basadas en AFC pueden ser utilizadas para encontrar vecinos (conceptos similares) de una forma eficiente y correcta. Una aproximación similar y reciente para el mismo tipo de problema con similares resultados se presenta en (Li y Murata, 2010). Estos trabajos muestran que la aplicación de técnicas basadas en AFC puede ser utilizada con éxito en los sistemas de recomendación colaborativos.

Nuestra propuesta para el módulo de pre-filtrado sigue esta línea, utilizando técnicas de AFC para mejorar los resultados. Así, en primer lugar, nuestro objetivo es desarrollar especificaciones contextuales más flexibles mediante la consideración de relaciones difusas en AFC. Esta

extensión fue introducida inicialmente en (Belohlavek, 1999), aunque con ciertos problemas relacionados con el desarrollo de nuevos métodos para inferir los conceptos y manejar implicaciones en relaciones difusas. Para resolverlos, se han aplicado los resultados teóricos presentados en (Belohlavek et al., 2012), desarrollando una lógica difusa correcta y completa para el manejo de dependencias funcionales, que constituye un marco de trabajo adecuado para la gestión eficiente de las implicaciones.

6.2.2. Marco teórico

Como se indicó anteriormente, nuestro objetivo es añadir a las especificaciones contextuales y de preferencias de usuarios un grado de imprecisión y de incertidumbre por medio de la lógica difusa. Desde que Lofti Zadeh introdujo la teoría de conjuntos difusos (Zadeh, 1996), la aproximación más habitual es reemplazar el conjunto de valores de verdad $\{0,1\}$ por un retículo residuado arbitrario. Nuestra propuesta utiliza una extensión de los retículos residuados, específicamente $([0,1], \wedge, \vee, 0, 1, \otimes, \rightarrow, *, \backslash)$ en el cual el intervalo $[0,1]$ está dotado de las siguientes operaciones:

- El ínfimo (denotado por \wedge) que juega el rol de cuantificador universal.
- El supremo (denotado por \vee) que juega el rol de cuantificador existencial.
- Una t-norma arbitraria continua por la izquierda (denotada por \otimes) como la conjunción.
- El residuo definido como:

$$a \rightarrow b = \sup\{x \in [0,1] | x \otimes a \leq b\}$$

- Operador de verdad reforzada, en inglés “*truth-stressing hedge*” (denotado por *)
- Diferencia (denotada por \)

Las t-normas más utilizadas son el producto estándar, producto de Lukasiewicz y el producto de Gödel. Los operadores de verdad reforzada (Hájek, 2001) conforman un tipo de funciones unarias de verdad que capturan la semántica de la noción “muy cierta”. Los dos casos extremos son la identidad y la globalización, cumpliéndose que:

$$1^* = 1 \text{ y } x^* = 0, \forall 1 \neq x \in L$$

La operación diferencia es dada por:

$$x \setminus y = x \text{ si } y < x, x \setminus y = 0 \text{ en otro caso.}$$

Un conjunto difuso sobre un universo U se define como una función $A: U \rightarrow [0,1]$ y el conjunto de operadores definidos como sigue:

Dados $A, B: U \rightarrow [0,1], \forall u \in U$

$$(A \cup B)(u) = A(u) \vee B(u)$$

$$(A \cap B)(u) = A(u) \wedge B(u)$$

$$(A \otimes B)(u) = A(u) \otimes B(u)$$

$$(A \rightarrow B)(u) = A(u) \rightarrow B(u)$$

$$(A \setminus B)(u) = A(u) \setminus B(u)$$

$$A^*(u) = (A(u))^*$$

Además, \emptyset y U son los conjuntos difusos en los cuales:

$$\forall u \in U, \emptyset(u) = 0 \text{ y } U(u) = 1$$

La inclusión de conjuntos puede ser extendida como sigue: para $A, B: U \rightarrow [0,1]$, el grado en el cual A es un subconjunto de B es

$$S(A, B) = \bigwedge_{u \in U} (A(u) \rightarrow B(u))$$

En el caso que $S(A, B)=1$ diremos que $A \subseteq B$ y en este caso

$$A(u) \leq B(u), \forall u \in U$$

El modelo propuesto trabaja con conjuntos difusos finitos, esto es, conjuntos difusos en los cuales, a lo sumo, un número finito de elementos tiene valores no nulos. En la notación que utilizaremos, los elementos con valores cero, no aparecen y el grado 1 se omite. Así, por ejemplo:

$$A = \{b/_{0.4}, d/_{0.1}, f\}$$

denota que

$$A(b) = 0.4, A(d) = 0.1, A(f) = 1, A(x) = 0 \text{ en otro caso}$$

Como hemos indicado anteriormente, la información del sistema de recomendación la vamos a organizar utilizando la extensión difusa del AFC, introducida en (Belohlavek, 1999) que puede ser considerada como la tendencia actual en esta área.

El punto de comienzo en el AFC difuso es la relación difusa que captura el grado que un objeto posee un atributo dado. Más específicamente, dado un conjunto finito de objetos X y un conjunto de atributos Y, la técnica AFC difusa extrae conocimiento a partir de la relación difusa

$$I: X \times Y \rightarrow [0,1]$$

donde

$$I(x, y) = \vartheta$$

sabiendo que ϑ , es el grado en el cual el objeto x posee el atributo y.

Normalmente, la relación difusa I es mostrada en una tabla, en la cual las filas representan los objetos y las columnas corresponden a los atributos. Así, la posición $[x,y]$ de la tabla representa el grado $I(x,y)$.

Una información importante que puede ser extraída a partir de las relaciones difusas es obtenida en términos de implicaciones de atributos. Existirán fórmulas del tipo $A \Rightarrow B$, donde A y B son conjuntos difusos de atributos. El grado, en el cual esta implicación de atributos es satisfecha por medio de una relación difusa I, viene dada por:

$$\|A \Rightarrow B\|_I = \bigwedge_{x \in X} (S(A, I_x)^* \rightarrow S(B, I_x))$$

donde I_x denota el conjunto difuso en el que $I_x(y) = I(x, y) \forall y \in Y$.

Así, por ejemplo:

$$\{b/_{0.2}, d\} \Rightarrow \{c/_{0.8}\}$$

significa que para todo objeto que tenga el atributo b con un grado al menos de 0.2 y el atributo d con un grado 1, tiene un atributo c con un grado de al menos 0.8.

Es importante destacar que tanto la parte izquierda como derecha de la implicación (los conjuntos A y B) pueden estar vacías. Si el conjunto B está vacío, la implicación captura una información que siempre es

válida, no teniendo que ser considerado en el proceso de inferencia. Sin embargo, si el conjunto A está vacío, la implicación suministra una información relevante, en particular en la aplicación sobre la que estamos trabajando. Por ejemplo, la implicación $\emptyset \Rightarrow \{c/_{0.8}\}$, es interpretada como que el atributo c debe tener al menos el grado 0.8.

Dada una relación difusa $I, A \xrightarrow{\vartheta} B$, denota que la implicación $A \Rightarrow B$ es cierta con grado al menos ϑ , lo que es equivalente a asegurar que

$$A \Rightarrow \vartheta \otimes B$$

es satisfecha con grado 1 (Belohlavek et al., 2012). Por tanto, usaremos $\vartheta \otimes B$ para denotar el conjunto difuso $(\vartheta \otimes B)(y) = \vartheta \otimes B(y) \forall y \in Y$. Los resultados citados aseguran que el usuario puede especificar implicaciones con grados que posteriormente se pueden traducir a implicaciones sin grados, que pueden ser tratadas de modo más simple mediante métodos automatizados.

6.2.3. Una lógica para el manejo de implicaciones de atributos difusos

Para el manejo de la información en términos de implicaciones, se requiere un método automático con objeto de inferir nueva información a partir de un conjunto de implicaciones. Aunque en la literatura existente podemos encontrar métodos de sistemas axiomáticos, no han sido desarrollados para el diseño de tales métodos automáticos. En (Belohlavek, 2012) los autores presentan la lógica de simplificación de atributos FSL, proporcionando por primera vez un sistema axiomático correcto y completo para el razonamiento con implicaciones mediante

un método de deducción automática. El sistema consiste en tres reglas de deducción:

Axioma:

$$\vdash AB \Rightarrow A$$

Simplificación:

$$A \Rightarrow B, C \Rightarrow D \vdash A(C \setminus B) \Rightarrow D$$

Multiplicación:

$$A \Rightarrow B \vdash \vartheta^* \otimes A \Rightarrow \vartheta^* \otimes B$$

donde A , B , C y D son conjuntos difusos y $\vartheta \in [0,1]$. En adelante usaremos las convención de escribir AB en lugar de $A \cup B$.

La característica más relevante de este sistema axiomático radica en que las reglas de inferencia pueden ser vistas como reglas de equivalencia, que nos permiten eliminar (simplificar) información redundante. Esta simplificación es aplicada tanto en las implicaciones como en los atributos dentro de las implicaciones. Esto es especialmente apropiado en entornos heterogéneos, donde diferentes usuarios suministran varias implicaciones, lo que hace muy probable que se generen especificaciones con un alto grado de redundancia. Estas equivalencias son las siguientes:

1. $\{A \Rightarrow B\} \equiv \{A \Rightarrow B \setminus A\}$
2. $\{A \Rightarrow B, A \Rightarrow C\} \equiv \{A \Rightarrow BC\}$
3. $\{A \Rightarrow B, C \Rightarrow D\} \equiv \{A \Rightarrow B, (C \setminus B) \Rightarrow D \setminus B\}$ cuando $A \subseteq C$

donde A, B, C y D son conjuntos difusos.

6.2.4. Aplicación de FSL a los sistemas de recomendación contextuales

En los apartados anteriores se han presentado los fundamentos teóricos que serán utilizados para incorporar el contexto en un sistema de recomendación. Como ya hemos indicado, nuestra propuesta utiliza implicaciones de atributos difusos y se basa en las siguientes características:

- a) La lógica difusa y el análisis formal de conceptos difuso multivaluado han demostrado ser una buena solución para poder especificar y razonar con incertidumbre. Ganter y Wille (Ganter y Wille, 1999), proponen transformar un contexto multivaluado (contextos con atributos que pueden tener un conjunto de valores) en un contexto de valores simples utilizando lo que denominan el escalado conceptual que en nuestro caso tomará valores en el intervalo [0,1].
- b) Se propone una combinación unificada de las técnicas de razonamiento basadas en contexto como un proceso de prefiltrado dentro de un marco de trabajo de recomendación.
- c) Se asocia cada contexto con un conjunto de implicaciones difusas definidas sobre las características de los ítems. Estas implicaciones han de ser diseñadas por expertos del dominio del problema (en nuestro caso, técnicos de turismo del destino).
- d) El usuario introduce en el sistema todas las características de su contexto, para que los métodos de razonamiento depuren el

conjunto de todas las implicaciones asociadas, y así obtener un conjunto de implicaciones difusas equivalente y más simple.

- e) El conjunto final de implicaciones es utilizado para validar y por tanto podar, los ítems que pueden pasar a los motores de recomendación. Así, el conjunto original de ítems es pre-filtrado y sólo este subconjunto filtrado formará parte de los ítems de entrada para los diferentes motores de recomendación.

A continuación, vamos a detallar mediante un ejemplo cómo se representa la información contextual mediante implicaciones difusas. Los diferentes puntos turísticos son representados por medio de un conjunto de atributos que describen sus características (si es barato o caro, si el ambiente es romántico o alegre, etc.). De esta forma se resuelve el problema de que los datos en el mundo real son frecuentemente complejos y difíciles de ser calificados con un dominio binario, sin pérdida de información. Por ejemplo, un restaurante puede tener un bonito jardín con varias mesas donde los niños pueden divertirse, pero también un salón con bastante intimidad y con música tranquila ideal para una cena romántica. Nuestra propuesta para almacenar la información de estas características de los ítems es por medio de relaciones difusas, como por ejemplo la que se muestran en las tablas 6.2 y 6.3. Así, cada fila corresponde a un objeto y cada columna a un atributo.

Restaurante est\'andar	i_1
Restaurante Estrella Michel\'in	i_2
Burger	i_3
Tapas Bar	i_4
Pizzeria	i_5
Marisquer\'ia de lujo en playa	i_6

Tabla 6.2. Lista de ítems

	Diseño		Ambiente			Precio			Servicios		
	Espacio abierto	Espacio cerrado	Tranquilo	Animado	Pintoresco	Barato	Moderado	Caro	Air.Ac	Vistas	Terraza
i_1	0.3	0.8	0.8	0.5	0.2	0.7	0.3	0.3	0.3	0.3	0.3
i_2	0.1	0.8	0.9	0.2	0.1	0	0.1	0.9	0.9	0.5	0.1
i_3	0.3	0.8	0.3	0.8	0.1	0.9	0.3	0.1	0.8	0.1	0.4
i_4	0.3	0.8	0.2	0.8	0.9	0.9	0.5	0.1	0.5	0.1	0.1
i_5	0.1	0.9	0.3	0.8	0.7	0.9	0.5	0.3	0.8	0.3	0.5
i_6	0.9	0.2	0.3	0.8	0.8	0.5	0.7	0.8	0.3	0.9	0.9

Tabla 6.3. Representación en AFC de los ítems y sus características

En este ejemplo hemos considerado un grupo de restaurantes de un destino turístico con algunos atributos que lo describen (diseño, ambiente, precio y comodidades). Cada atributo tiene un conjunto finito de posibles valores y un experto del destino, que maneja el sistema, suministrará los grados de los diferentes valores. Así, se obtiene una tabla de restaurantes con sus grados para obtener la tabla difusa (Tabla 6.3).

El contexto del sistema es representado por medio de un conjunto de dominios discretos $C = \{C_1, C_2, \dots, C_n\}$. Cada dominio se asocia con una dimensión del contexto (por ejemplo tiempo, compañía, horario, etc.) que tiene, a su vez, un conjunto finito de valores $C_i = \{v_1^i, v_2^i, \dots, v_n^i\}$. Así, definimos el contexto del usuario

(que denominaremos *estado*) como una n-tupla de pares (valor del dominio, grado).

Por ejemplo, supongamos que el contexto tiene tres dimensiones contextuales, tiempo, compañía y horario, con los siguientes dominios:

$$\text{Tiempo} = \{\text{caluroso}, \text{nublado}, \text{lluvioso}\}$$

$$\text{Compañía} = \{\text{solo}, \text{amigos}, \text{pareja}, \text{familia}, \text{grupo}\}$$

$$\text{Horario} = \{\text{mañana}, \text{mediodia}, \text{tarde}, \text{noche}\}$$

El sistema debe permitir al usuario especificar su contexto (el propio turista tiene que introducir algunos valores, pero otros como por ejemplo el tiempo podrían ser suministrados automáticamente por el sistema). Supongamos que un turista que desea una recomendación está representado por el siguiente estado:

$$\{(caluroso, 0.8), (tarde, 0.8), (familia, 0.7)\}$$

Se define un segmento de contexto como un valor específico de un dominio y su grado asociado. Nuestra propuesta proporciona un marco de trabajo donde cada segmento contextual es asociado a un conjunto de implicaciones difusas. Como hemos indicado anteriormente, las implicaciones pueden ser etiquetadas con un grado para expresar la certeza de la propia implicación. Este grado del segmento contextual es heredado por todas las implicaciones.

Las implicaciones asociadas a cada segmento contextual son introducidas como sigue (obsérvese que el grado del contexto es heredado por cada una de las implicaciones):

Segmento contextual:

*caluroso/*_{0.8}

Implicaciones:

*Caro/*_{0.8}, *Espacio cerrado/*_{0.8} $\xrightarrow{0.8}$ *Aire ac./*_{0.8}, *Vistas/*_{0.9}, *Pintoresco/*_{0.2}

*Espacio abierto/*_{0.8} $\xrightarrow{0.8}$ *Barato/*_{0.7}

Segmento contextual:

*tarde/*_{0.8}

Implicaciones:

*Espacio abierto/*_{0.8} $\xrightarrow{0.8}$ *Terraza/*_{0.6}, *Barato/*_{0.9}

Segmento contextual:

*familia/*_{0.7}

Implicaciones:

$\emptyset \xrightarrow{0.7} \text{barato/}_{0.6}$

*Espacio cerrado/*_{0.8} $\xrightarrow{0.7}$ *Aire ac./*_{0.9}

La implicación $\emptyset \xrightarrow{0.7} \text{barato/}_{0.6}$ indica que si el usuario está acompañado por su familia con un grado 0.7, entonces el restaurante debe ser barato con un grado al menos 0.6.

Cuando un sistema maneja una gran cantidad de información, es importante disponer una forma automática para analizar y extraer la

información relevante para de esta forma reducir los costes computacionales. En el modelo propuesto se usarán métodos de deducción automática desarrollada sobre la lógica FSL para depurar la especificación del contexto y obtener un conjunto canónico de implicaciones.

El siguiente ejemplo, ilustra los beneficios del uso de FSL para obtener un conjunto equivalente y más simple de implicaciones que represente la información del contexto.

A partir de la especificación del ejemplo anterior, si tenemos el contexto suministrado por medio del usuario como:

$$\{caluroso/_{0.8}, tarde/_{0.8}, familia/_{0.7}\}$$

A partir de él, se construye el conjunto de implicaciones incluyendo todas las implicaciones anteriores en un conjunto unificado:

$$Caro/_{0.8}, Espacio cerrado/_{0.8} \xrightarrow{0.8} Aire ac./_{0.8}, Vistas/_{0.9}, Pintoresco/_{0.2}$$

$$Espacio abierto/_{0.8} \xrightarrow{0.8} Barato/_{0.7}$$

$$Espacio abierto/_{0.8} \xrightarrow{0.8} Terraza/_{0.6}, Barato/_{0.9}$$

$$\emptyset \xrightarrow{0.7} barato/_{0.6}$$

$$Espacio cerrado/_{0.8} \xrightarrow{0.7} Aire ac./_{0.9}$$

Usando las reglas de inferencia de FSL, podemos eliminar información redundante y obtener un conjunto más simple y equivalente de implicaciones. A continuación exponemos las diferentes simplificaciones:

Fase 1:

Caro/0.8, Espacio cerrado/0.8 ⇒ Aire ac./0.6, Vistas/0.7, Pintoresco/0.1

Espacio abierto/0.8 ⇒ Barato/0.5

Espacio abierto/0.8 ⇒ Terraza/0.4, Barato/0.7

$\emptyset \Rightarrow barato/0.4$

Espacio cerrado/0.8 ⇒ Aire ac./0.6

Fase 2:

Caro/0.8, Espacio cerrado/0.8 ⇒ Aire ac./0.6, Vistas/0.7, Pintoresco/0.1

Espacio abierto/0.8 ⇒ Barato/0.5

Espacio abierto/0.8 ⇒ Terraza/0.4, Barato/0.7

$\emptyset \Rightarrow barato/0.4$

Espacio cerrado/0.8 ⇒ Aire ac./0.6

Fase 3:

Caro/0.8, Espacio cerrado/0.8 ⇒ Vistas/0.7, Pintoresco/0.1

Espacio abierto/0.8 $\overline{\Rightarrow}$ Barato/0.5

Espacio abierto/0.8 ⇒ Terraza/0.4, Barato/0.7

$\emptyset \Rightarrow barato/0.4$

Espacio cerrado/0.8 ⇒ Aire ac./0.6

Fase 4:

$$\text{Caro}/_{0.8}, \text{Espacio cerrado}/_{0.8} \Rightarrow \text{Vistas}/_{0.7}, \text{Pintoresco}/_{0.1}$$

$$\text{Espacio abierto}/_{0.8} \Rightarrow \text{Terraza}/_{0.4}, \text{Barato}/_{0.7}$$

$$\emptyset \Rightarrow \text{barato}/_{0.4}$$

$$\text{Espacio cerrado}/_{0.8} \Rightarrow \text{Aire ac.}/_{0.6}$$

Es necesario destacar que el algoritmo de eliminación de redundancia tiene una complejidad cuadrática respecto al número de implicaciones. Este número es mucho menor que el total de ítems (normalmente varios miles) de un destino turístico.

Nuestro modelo hace uso de la información asociada al contexto del usuario, generando un conjunto unificado y depurado de implicaciones que deben forzosamente cumplir los ítems que superan el proceso de pre-filtrado. Para cada ítem de la tabla de AFC, se debe validar el conjunto de implicaciones, eliminando todos los ítems que no lo satisfagan. La complejidad de esta última etapa es $O(n)$ donde n es el número de ítems. De esta manera, el proceso de pre-filtrado contextual que hemos desarrollado tiene complejidad lineal.

Siguiendo el ejemplo en el que el contexto del usuario es la tarde de un día caluroso y viajando con la familia, nuestro proceso de pre-filtrado contextual reduce la lista de ítems del ejemplo solamente a i_3 e i_5 (Burger y Pizzeria). Más detalladamente:

- i_1 (Restaurante estándar) e i_4 (Tapas Bar) no satisfacen la implicación

$$Espacio\ cerrado/_{0.8} \Rightarrow Aire\ ac./_{0.6}$$

- i_2 (Restaurante con estrellas Michelín) no satisface la implicación

$$Caro/_{0.8}, Espacio\ cerrado/_{0.8} \Rightarrow Vistas/_{0.7}, Pintoresco/_{0.1}$$

- i_6 (Marisquería de lujo en la playa) no satisface la implicación

$$\emptyset \Rightarrow barato/_{0.4}$$

De esta forma, el número de ítems que deben ser manejados por el sistema de recomendación se ve significativamente reducido.

6.2.5. Consideraciones finales

El método propuesto (PFCBI) filtrará aquellos ítems que no cumplan las implicaciones contextuales, a partir del contexto expresado por el turista en ese momento. Es una importante mejora, ya que se conseguirá reducir drásticamente el número de ítems que deben ser entrada de los motores de recomendación que exponemos en las siguientes secciones.

Se ha desarrollado una aplicación web que implementa el proceso de pre-filtrado contextual. Esto ha permitido realizar diferentes pruebas sobre un conjunto de entrada de 1240 ítems a los que se les han aplicado diferentes situaciones contextuales (Figura 6.3).

The screenshot shows a web-based application titled "POI Recommender" from the University of Málaga. The interface includes a navigation bar with tabs for "Database", "Filtering level", "Simplification", "Filter Settings", and "Results". The "Results" tab is active, displaying a table of 15 items. The columns are "Name", "City", "Province", and "ZIP", followed by a column for "Result" which contains green "true" or red "false" status indicators. Most items have a green "true" status, while three items (PLAZA DE ESPAÑA, PLAZA GUACÍN, and PLAZA NUESTRA SEÑORA DEL ROSARIO) have a red "false" status. Below the table is a pagination control with links from < 1 2 3 4 5 6 7 8 9 10 11 >. A "Back" button is at the bottom left, and a copyright notice "©2013 SICUMA - Sistemas de Información Colaborativos de la Universidad de Málaga" is at the bottom right.

Figura 6.3. Herramienta diseñada para simular el motor PFCBI

Como se puede observar en la figura 6.4 el porcentaje de ítems que se descartan es bastante significativo en la mayoría de los casos.

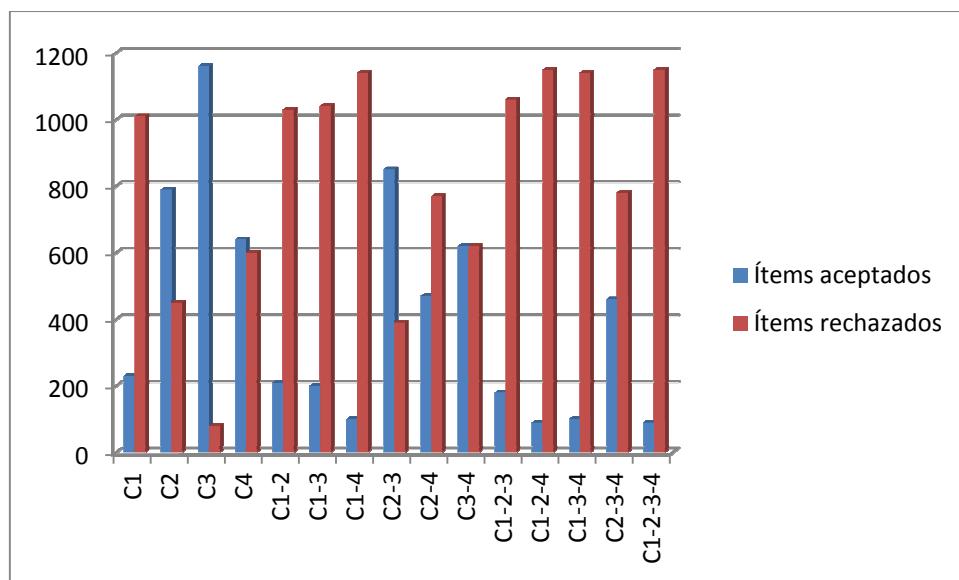


Figura 6.4. Comparativa ítems pre-filtrados y rechazados por el motor PFCBI

En particular, hemos denominado C1, C2, C3 y C4 a cuatro situaciones contextuales diferentes. Así, C1-3-4 debe interpretarse como que se presentan las situaciones contextuales 1, 3 y 4.

Es evidente que es difícil hacer que el usuario exprese explícitamente todas las características contextuales, pero aún con un subconjunto de ellas será posible realizar un pre-filtrado previo. Como se deduce de los resultados ilustrados en la figura 6.4 a mayor conocimiento de la situación contextual, aumenta el número de ítems rechazados.

Como se verá en los siguientes apartados, en los motores de recomendación definidos en este trabajo, se vuelve a hacer uso de los atributos contextuales, además de permitir otros atributos que puedan no estar presentes en las implicaciones del pre-filtrado difuso.

Finalmente, se ha mostrado como la utilización de lógica difusa y el análisis de formal conceptos son una herramienta sólida para combinar información de contexto y sistemas de recomendación, constituyendo un proceso inteligente de pre-filtrado con un coste lineal.

6.3. Motor de recomendación basado en contenido sin memoria (SPP^u)

En esta sección se describe el motor de recomendación basado en contenido sin memoria propuesto en nuestro modelo. Este motor de recomendación trabaja sobre las características que describen los diferentes ítems susceptibles de recomendación. Se comenzará realizando un resumen del funcionamiento básico, profundizando en aspectos

relacionados con las características y aspectos formales de la propuesta para la realización del proceso de recomendación. Además, se aporta un ejemplo de dicho proceso.

Para que el modelo tenga utilidad para un destino turístico, las descripciones y valoraciones de las características de las diferentes actividades o puntos turísticos tienen que ser realizadas por personal cualificado y experto. Todo ítem, que pueda ser recomendado por el motor, tiene que ser descrito adecuadamente por expertos del destino. Por ejemplo, un técnico de turismo debería informar sobre servicio, precios, calidad de comida, comodidades, etc., de los establecimientos de restauración.

Los motores de recomendación basados en contenido se encargan de buscar ítems que tengan descripciones similares a las necesidades planteadas por el turista. El principal problema que se puede encontrar es que las descripciones de las características de un producto turístico realizadas por un experto pueden ser bastante diferentes a las características descritas en las necesidades de un turista, ya que el experto puede tener mayor capacidad para matizar la descripción.

En muchos casos las propiedades o atributos de un producto pueden ser descritas mediante un valor numérico, pero en otros casos, puede ser una solución muy adecuada utilizar una descripción cualitativa mediante la elección de un valor que la mayor parte de las veces pertenece a un dominio lingüístico. Debido a la diversidad de características es adecuado utilizar diferentes escalas, es decir, se deben

suministrar mecanismos que hagan posible presentar a los turistas un contexto flexible que les permita utilizar escalas lingüísticas adecuadas.

6.3.1. Descripción

Es conveniente recordar que previamente a la ejecución del motor de recomendación propuesto en esta sección, se debe realizar la fase de pre-filtrado contextual basado en implicaciones (PFCBI) descrita en el apartado anterior. En esta fase, partiendo de información contextual del usuario y mediante implicaciones difusas, se reducirá considerablemente el número de ítems de entrada al motor de recomendación, con objeto de mejorar los tiempos de ejecución de éste.

Para usar este motor de recomendación, el turista deberá proporcionar explícitamente información sobre las características más importantes de aquellos productos que le gustaría visitar. Además, también deberá suministrar información sobre su tipología de turista (obtenida del estereotipo determinado por los descriptores turísticos y demográficos definidos en el capítulo anterior⁸⁸) e información contextual relevante. El motor de recomendación parte entonces de las necesidades⁸⁹ del turista, es decir, de qué le gustaría visitar, sin olvidar el tipo de turista y sus características.

Se trata, por tanto, de un “perfil actual” del turista con sus preferencias actuales. De esta forma, pueden expresarse necesidades que no son habituales, filtrando el propio turista el tipo de ítem que desea visitar, como podría ser, por ejemplo, restaurantes o museos.

⁸⁸ Definiciones 5.15, 5.16 y 5.17

⁸⁹ Definición 5.27

En la mayoría de los sistemas de recomendación basados en contenido (Schwab, Kobsa y Koychev, 2001; Mak, Koprinska y Poon, 2003; Bezerra y Carvalho, 2004; Pérez, 2008), encontramos que el mecanismo de obtención de la información del perfil del turista, parten de información histórica y no utilizan información contextual (Leiva, Guevara y Rossi, 2012). En cambio, en el motor que estamos describiendo, esa fuente de información histórica es sustituida por una necesidad actual especificada de forma explícita y por información complementaria sobre la tipología de turista y su contexto.

Además de la información de los turistas, también tendremos la información sobre los puntos y actividades turísticas. En concreto, se utilizarán las descripciones de los productos (valor de diferentes características de los ítems, así como información sobre el tipo de turistas que suele visitar esos puntos o actividades). Esta información, como se ha indicado anteriormente debe ser introducida por expertos en el dominio del destino turístico.

En resumen, este motor de recomendación, partiendo del “perfil actual” del turista y de información sobre los puntos de interés, realizará un proceso de filtrado en el que se calculará la similitud entre el “perfil actual” del turista y los diferentes ítems a recomendar. En este punto es importante recordar la posibilidad de que las descripciones de las características puedan estar expresadas en diferentes escalas, lo que hará necesario definir operaciones para poder comparar los valores de las características de los productos y del “perfil actual” del turista. Más detalladamente, estas operaciones tienen como objetivo convertir los

valores de las diferentes escalas a valores entre 0 y 1. Una vez aplicados estos cálculos, el sistema recomendará, obviamente, aquellos puntos que tengan un grado de similitud mayor entre sus características y las necesidades del turista.

6.3.2. Tipos de escalas

Se pueden encontrar diferentes situaciones a la hora de realizar las valoraciones por parte de un experto y por parte de un turista. Estas situaciones son:

- a) El experto y el turista utilizan la misma escala para valorar una característica, tratándose de un valor numérico.
- b) El experto y el turista utilizan la misma escala para valorar una característica, tratándose de un valor cualitativo.
- c) El experto y el turista utilizan diferentes escalas para valorar una característica y es un valor numérico.
- d) El experto y el turista utilizan diferentes escalas para valorar una característica y es un valor cualitativo.

El principal problema puede darse en los casos en los que se utilicen valoraciones cualitativas. El conocimiento de una característica o atributo por parte de un experto y de un turista puede ser diferente, por lo que, a veces, pueden utilizar escalas lingüísticas distintas para expresar el valor de dicho atributo. En este caso, se deben realizar procesos que permitan la comparación entre valores pertenecientes a escalas lingüísticas diferentes. Para ello, en este trabajo optamos por representar sus semánticas mediante valores difusos.

6.3.3. Preliminares

Cuando la información que se usa para la toma de decisiones es imprecisa, resulta más adecuado utilizar información cualitativa que cuantitativa. En estas situaciones, la utilización del enfoque lingüístico difuso resulta una herramienta bastante útil que proporciona buenos resultados (Arfi, 2005). En el caso que nos ocupa, las características contextuales de los turistas o las características de los productos turísticos que van a ser recomendados, pueden ser expresadas mediante el uso de un modelado lingüístico.

El enfoque lingüístico difuso proporciona un modelo de representación de información mediante el uso de variables lingüísticas, cuya principal diferencia con las variables numéricas radica en que sus valores representan aspectos cualitativos. Si una variable puede tomar sus valores partiendo de palabras en un lenguaje natural, ésta es llamada variable lingüística (Xu, 2004; Galindo, Urrutia y Piattini, 2006; Ma et al., 2007; Martinez et al., 2008). Por tanto, en el enfoque difuso una variable lingüística se caracteriza por un valor sintáctico y un valor semántico. El valor sintáctico viene determinado por una palabra perteneciente a un conjunto de términos lingüísticos, mientras que el significado viene determinado por un conjunto difuso en el universo del discurso en el cual se define la variable. A continuación se exponen brevemente algunas definiciones relacionadas con el enfoque lingüístico difuso (Zadeh, 1996; Herrera, Herrera-Viedma y Martínez, 2000; Xu, 2004; Arfi, 2005; Martinez et al., 2008; Porcel y Herrera-

Viedma, 2010; Rodriguez, Martínez Y Herrera, 2012) asociándolas al ámbito turístico que nos ocupa.

Definición 6.1: Una variable lingüística es caracterizada por una tupla $(X, T(X), U, M)$, donde X es el nombre de la variable lingüística, $T(X)$ es el conjunto de valores o etiquetas lingüísticas que puede tomar X , U el dominio físico real en el cual la variable lingüística X toma sus valores cuantitativos (nítidos) y M es una regla semántica que relaciona cada valor lingüístico en T con un conjunto difuso en U ⁹⁰.

Para que los turistas que visitan el destino y los expertos del mismo puedan valorar con facilidad sus opiniones o preferencias, es necesario disponer de un conjunto adecuado de descriptores lingüísticos, así como de un conjunto de valores lingüísticos disponibles para expresar esta información. Con el objetivo de conseguir generar los términos o valores lingüísticos, utilizaremos el enfoque basado en términos con una estructura ordenada, que consiste en asignar un conjunto de términos a una escala con un orden total definido (Xu, 2004).

Para ilustrar la definición anterior, vamos a presentar un ejemplo en el que se utiliza un conjunto de cinco etiquetas o valores, para describir si un turista considera que realiza habitualmente turismo cultural:

$$T(X) = \{J, A, N, B, T\}$$

$$v_0 = J = \text{Jamás}$$

$$v_1 = A = \text{Algunas veces}$$

⁹⁰ Un conjunto difuso podemos definirlo informalmente como una colección de objetos con valores de pertenencia entre 0 y 1. Se define más formalmente en la definición 6.2

$v_2 = N = \text{Normalmente si}$

$v_3 = B = \text{Con frecuencia}$

$v_4 = T = \text{Siempre}$

Se define la relación de orden $v_i < v_j \Leftrightarrow i < j$

Así si el conjunto de etiquetas está definido por $\{v_0, v_1, \dots, v_g\}$ los valores lingüísticos deben satisfacer las siguientes condiciones (Martínez et al., 2008):

$$\text{Cardinalidad}(T(X)) = g + 1$$

donde Neg es un operador de negación definido como sigue:

$$\text{Neg}(v_i) = v_j, \text{ donde } j = g - i,$$

$$\min(v_i, v_j) = v_i \Leftrightarrow v_i \leq v_j$$

$$\max(v_i, v_j) = v_j \Leftrightarrow v_i \leq v_j$$

Definición 6.2: Se define un conjunto difuso como aquel cuyos elementos poseen un grado de pertenencia, expresado mediante un valor entre 0 y 1. Éste vendrá determinado por una función de pertenencia del conjunto, que se define a continuación.

En un conjunto clásico, todos sus elementos tendrían un grado de pertenencia igual a 1.

Definición 6.3: La función de pertenencia de un conjunto difuso A, denotada F_A , es la función que determina para cada elemento de un

universo dado (X) su grado de pertenencia al conjunto A. Se define como

$$F_A: X \rightarrow [0,1] \text{ donde } F_A(x) = r$$

donde r es el grado en que x pertenece a A, siendo r , obviamente, un valor numérico perteneciente al intervalo $[0,1]$. El valor 1 indica máxima pertenencia y el valor 0 indica nula pertenencia.

Para definir la semántica de un conjunto de términos lingüísticos, suelen utilizarse las funciones de pertenencia, que emplean valores difusos en el intervalo $[0,1]$, donde cada número difuso está definido por la función de pertenencia. El valor 1 indica máxima pertenencia y el valor 0 indica nula pertenencia.

Las funciones de pertenencia permiten representar de forma gráfica un conjunto difuso. En el eje de abscisas se representa el universo en discurso, mientras que en el eje de ordenadas se sitúan los grados de pertenencia en el intervalo $[0,1]$. Aunque en principio, para definir un conjunto difuso se podría utilizar cualquier función, existen algunos tipos especiales que por su simplicidad matemáticas son muy utilizados: triangulares, trapezoidales, gaussianas y pseudo-exponenciales, (Galindo, Urrutia y Piattini, 2006):

a) Funciones triangulares. Este tipo de función se definen por:

$$F_A(x) = \begin{cases} 0 & , \quad x \leq a \\ \frac{x-a}{m-a}, & a < x \leq m \\ \frac{b-x}{b-m}, & m < x < b \\ 0 & , \quad x \geq b \end{cases}$$

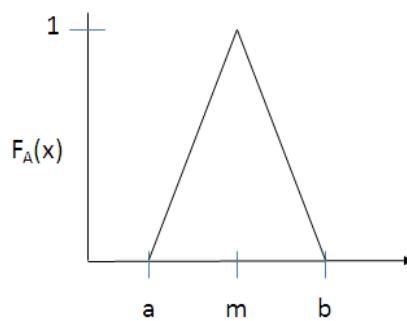


Figura 6.5. Función de pertenencia triangular

Las funciones de pertenencia triangulares simétricas suelen utilizar una notación basada en una tupla (a,m,b) denominada como número difuso (Nasseri, 2008), donde m es el valor donde la función de pertenencia toma el valor 1 (también conocido como valor modal), a es el extremo izquierdo o límite inferior de la función de pertenencia y b es el extremo derecho o límite superior de la función de pertenencia (Bojadziev y Bojadziev, 1995).

b) Funciones trapezoidales. Este tipo de función se definen por:

$$F_A(x) = \begin{cases} 0 & , \quad x \leq a \text{ o } x > b \\ \frac{x-a}{b-a}, & a < x < b \\ \frac{d-x}{d-c}, & c < x < d \\ 1 & , \quad b < x < c \end{cases}$$

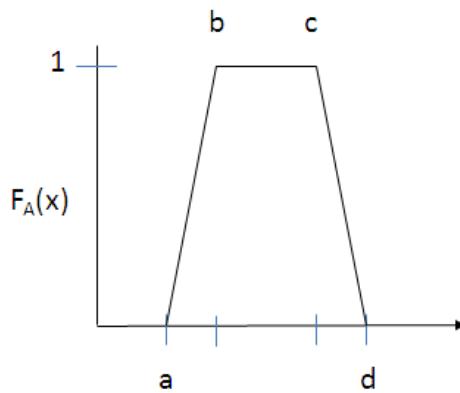


Figura 6.6. Función de pertenencia trapezoidal

- c) Funciones gaussianas. Quedan determinadas por su valor medio m y el valor $k > 0$. Una función de este tipo queda definida por:

$$F_A(x) = e^{-k(x-m)^2}$$

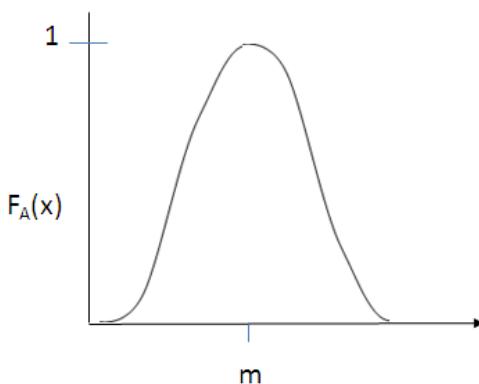


Figura 6.7. Función de pertenencia gaussiana

Cuanto mayor es el valor de k , más estrecha es la “campana”.

- d) Funciones pseudo-exponenciales. Quedan determinadas por su valor medio m y el valor $k > 1$. Una función de este tipo queda determinada por:

$$F_A(x) = \frac{1}{1 + k(x - m)^2}$$

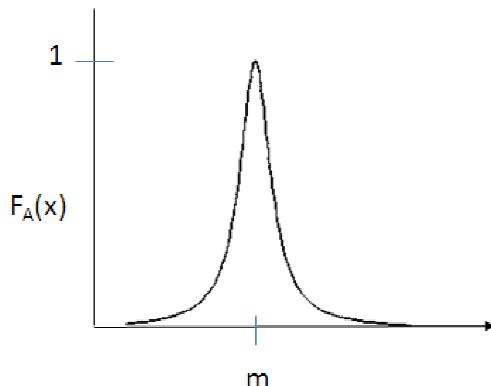


Figura 6.8. Función de pertenencia pseudo-exponencial

Cuanto mayor es el valor de k, el crecimiento es más rápido y más estrecha es la “campana”.

Por tanto, un conjunto difuso A en el universo X, puede representarse como un conjunto de pares ordenados de un elemento arbitrario x y su función de pertenencia del siguiente modo:

$$A = \{(x, F_A(x)), \text{ donde } x \in X \text{ y } F_A(x) \in [0,1]\}$$

Un aspecto a tener en cuenta al modelar lingüísticamente la información es la granularidad de la incertidumbre, es decir, el número de valores considerados al definir el conjunto de términos lingüísticos (Martínez et al., 2008). En el caso que nos ocupa la utilización de modelos multigranulares se debe a que el grado de conocimiento se va a aplicar a diferentes conceptos: categorización de turista, atributos contextuales y características de productos turísticos. Por tanto, es necesario utilizar escalas lingüísticas con diferente número de etiquetas. En la siguiente figura podemos observar varios ejemplos de escalas lingüísticas que utilizan función de pertenencia triangular. Las diferentes escalas representadas permiten distinguir entre 3, 5 y 7 etiquetas.

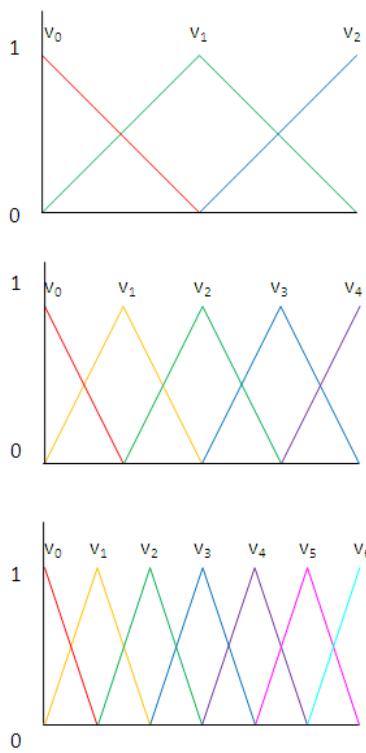


Figura 6.9. Escalas lingüísticas en un contexto multigranular

A continuación vamos a representar un ejemplo compuesto de cinco etiquetas con la que el usuario podría definir la característica “temperatura” (Figura 6.10). Por simplicidad, en el ejemplo se asume que el universo de discurso (representado en el eje de abscisas) es el intervalo $[0,1]$.

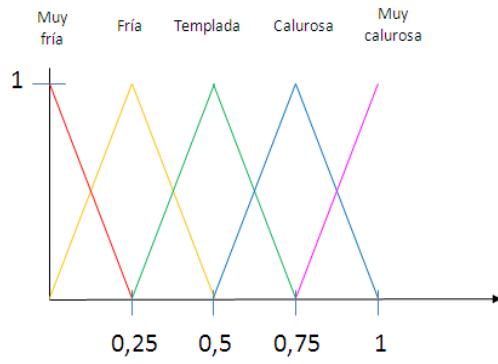


Figura 6.10. Ejemplo de varias funciones de pertenencia

Si se utiliza la notación mediante tuplas, la función de pertenencia de la etiqueta “Muy fría” (representada con la gráfica roja) se denotaría con la tupla (0, 0, 0.25), ya que su extremo izquierdo es 0, la función toma el valor 1 para el valor 0 del eje de abscisas, y su extremo derecho es 0,25. Análogamente, la función de pertenencia de la etiqueta “Fría” se denotaría con (0, 0.25, 0.5), la de “Templada” con (0.25, 0.5, 0.75), la de “Calurosa” con (0.5, 0.75, 1) y la de “Muy calurosa” con (0.75, 1, 1).

A continuación se presentan dos tipos de medidas de similitud para realizar comparaciones entre diferentes escalas lingüísticas: medida de posibilidad y medida de necesidad.

- La medida de posibilidad $\text{Poss}(A,B)$, mide en qué medida A y B se superponen, es decir, calcula la posibilidad de que el valor A sea igual al valor B (Dubois y Prade, 1983; Galindo, Urrutia y Piattini, 2006).

$$\text{Poss}(A,B) = \sup_{x \in X} \{\min(A(x), B(x))\}$$

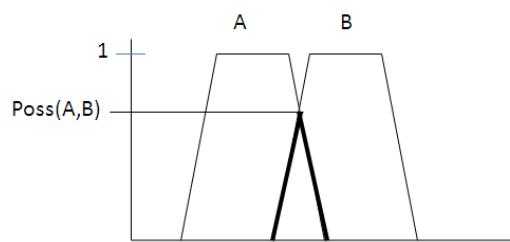


Figura 6.11. Medida de Posibilidad

- La medida de necesidad $\text{Nec}(A,B)$, mide la necesidad de A respecto de B, es decir, la necesidad de que el valor B sea igual al valor A, midiendo el grado con que B está incluido en A (Dubois y Prade, 1983; Galindo, Urrutia y Piattini, 2006).

$$\text{Nec}(A,B) = \inf_{x \in X} \{\max(A(x), 1 - B(x))\}$$

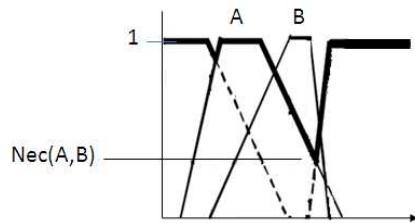


Figura 6.12. Medida de Necesidad

En nuestro caso, dado que habitualmente representamos la información mediante funciones de pertenencia triangulares, la medida de necesidad es demasiado restrictiva, pues siempre se obtendría un valor nulo. La medida de necesidad es más adecuada cuando se utilizan funciones de pertenencia trapezoidales.

6.3.4. Formalización

Sea I el conjunto de ítems (actividades turísticas o puntos de interés) que el sistema detecta a una distancia menor o igual a la indicada por el turista.

$$I = \{i_1, i_2, i_3, \dots, i_n\}$$

Sea U el conjunto de usuarios que utilizan el sistema propuesto.

$$U = \{u_1, u_2, u_3, \dots, u_m\}$$

Sea I_{τ} el conjunto índice de características contextuales {1, 2, ...t}, $I_{\mathcal{D}}$ el conjunto índice de descriptores turísticos y demográficos {1, 2,d}, e I_c el conjunto índice de características de un ítem {1, 2,c}.

A cada dominio D_i con $i \in I_{\tau}$ le añadiremos un nuevo elemento denominado ‘indiferente’ y denotado \perp que representará que no se dispone de información acerca de la característica c_i . Análogamente,

añadiremos el elemento \perp a cada dominio D_j con $j \in I_{\mathcal{D}}$ y para D_k con $k \in I_{\mathcal{C}}$.

Definición 6.4. Se define el concepto de descriptor del ítem i_k , denotado, $Ds(i_k)$, como una tupla formada por la concatenación de las listas de valores de atributos contextuales (t), valores de descriptores turísticos y demográficos (d) y de valores de características (c). Más formalmente:

$$Ds: I \rightarrow \prod_{i \in T} D_i \times \prod_{j \in D} D_j \times \prod_{k \in C} D_k$$

El descriptor $Ds(i_k)$ es por tanto

$$Ds(i_k) = (t_1^{ik}, t_2^{ik}, \dots, t_t^{ik}, d_1^{ik}, d_2^{ik}, \dots, d_d^{ik}, c_1^{ik}, c_2^{ik}, \dots, c_c^{ik})$$

Donde t_i^{ik} es el valor del atributo contextual i para el ítem i_k , con $i \in I_T$, d_j^{ik} es el valor del descriptor j para el ítem i_k con $j \in I_{\mathcal{D}}$ y c_m^{ik} representa la valoración de la característica m del ítem i_k con $m \in I_{\mathcal{C}}$.

Todos estos valores han de ser proporcionados por los expertos del destino.

Definición 6.5: Se denomina necesidad del turista u , denotado $N(u)$, como una tupla formada por la concatenación de las lista de valores de atributos contextuales (t), de valores de descriptores turísticos-demográficos (d) que lo definen y de valores de las características de los ítems que describen las preferencias “actuales” (p) del turista, Más formalmente:

$$N: U \rightarrow \prod_{i \in T} D_i \times \prod_{j \in D} D_j \times \prod_{k \in C} D_k$$

Las necesidad $N(u)$ es por tanto

$$N(u) = (t_1^u, t_2^u, \dots, t_t^u, d_1^u, d_2^u, \dots, d_d^u, p_1^u, p_2^u, \dots, p_c^u)$$

Donde t_i^u , es el valor del atributo contextual i para el usuario u con $i \in I_T$, d_j^u es el valor del descriptor j para el turista u con $j \in I_D$ y p_m^u representa la valoración de la característica m por parte del turista u con $m \in I_C$. Esta información es proporcionada por el turista.

Definición 6.6: Dado $j \in I_C$, se denomina escala experta del atributo j, denotada E^j , al conjunto de posibles valores o términos lingüísticos definidos para calificar el atributo j⁹¹.

$$E^j = \{e_1^j, e_2^j, \dots, e_z^j\}$$

Por tanto, la valoración de un atributo j por parte del experto del punto o actividad i_k , debe cumplir⁹²:

$$c_j^{ik} \in E^j, \quad j \in I_C$$

Definición 6.7: Dado $j \in I_C$, se denomina escala de usuario del atributo j, denotada EU^j , al conjunto de posibles valores o términos lingüísticos definidos para calificar el atributo j.

⁹¹ En la definición de variable lingüística (6.1) denominamos $T(x)$ al conjunto de valores o etiquetas lingüísticas que puede tomar una variable, pero debido a que en el sistema propuesto van a coexistir diferentes escalas, vamos a utilizar a partir de ahora esta nueva notación E^j .

⁹² De la misma forma se valoran los atributos contextuales y descriptores turísticos-demográficos,

$$EU^j = \{eu_1^j, eu_2^j, \dots, eu_z^j\}$$

Por tanto, la valoración del atributo j por parte del usuario u , debe cumplir⁹³:

$$p_j^u \in EU^j, \quad j \in I_c$$

6.3.5. Proceso de filtrado y recomendación

Como se detalló la sección 6.2, el sistema propuesto incluye un proceso de pre-filtrado contextual basado en implicaciones (PFCBI) mediante el cual se filtran los ítems detectados (I) por el sistema, eliminando aquellos que no satisfacen las implicaciones determinadas por el contexto del turista.

Una vez aplicado el proceso PFCBI, se propone aplicar un segundo proceso de filtrado basado en contenido sin memoria (FBSM) mediante el cual se determina qué ítems de los que han superado el filtro de la fase PFCBI deben ser recomendados teniendo en cuenta el perfil “actual” del turista. El objetivo de esta sección es describir el proceso FBSM.

En el proceso FBSM, a cada ítem resultante de la fase PFCBI se le aplica una medida de similitud que compare sus descriptores y características con las preferencias expresadas por el turista. A continuación se detallan los elementos necesarios para este proceso.

Definición 6.8: Se denomina similitud del atributo⁹⁴ k del ítem o actividad turística i y de la preferencia del usuario u en ese atributo k

⁹³ Igual que la nota anterior.

$(\text{sim}(k,i,u))$, al valor que mide cómo de parecidos son los valores lingüísticos asociados al ítem i y al usuario u . Al valor $\text{sim}(k,i,u)$, lo denotaremos a partir de ahora como sim_k^{i-u} . Este valor se calcula mediante la medida de posibilidad descrita en la sección anterior Poss(A,B).

Definición 6.9: Se denomina similitud contextual del atributo k del ítem i y del contexto actual del usuario u en ese atributo k ($\text{simc}(k,i,u)$), al valor que mide cómo de parecidos son los valores lingüísticos asociados al ítem i y al usuario u . Al valor $\text{simc}(k,i,u)$, lo denotaremos como simc_k^{i-u} . Este valor se calcula aplicando la medida de posibilidad descrita en la sección anterior Poss(A,B).

Definición 6.10: Se denomina similitud contextual del ítem i y del usuario u , a la lista obtenida a partir de la similitud de los atributos contextuales lingüísticos.

$$\text{Simc}(N(u), Ds(i)) = (\text{simc}_{t1}^{i-u}, \text{simc}_{t2}^{i-u}, \dots, \text{simc}_{tr}^{i-u})$$

considerando que el conjunto índice de características contextuales $I_{\tau} = \{1, 2, \dots, r\}$. En adelante denotaremos a $\text{Simc}(N(u), Ds(i))$ como Simc_i^u .

Definición 6.11: Se denomina similitud del ítem i y del usuario u , a la lista obtenida a partir de la comparación de los respectivos descriptores y características no contextuales.

$$\text{Sim}(N(u), Ds(i)) = (\text{sim}_{d1}^{i-u}, \text{sim}_{d2}^{i-u}, \dots, \text{sim}_{dq}^{i-u}, \text{sim}_{c1}^{i-u}, \text{sim}_{c2}^{i-u}, \dots, \text{sim}_{cm}^{i-u})$$

⁹⁴ Descriptor turístico-demográfico o característica

considerando que el conjunto índice de descriptores turísticos y demográficos $I_{\mathcal{D}} = \{1, 2, \dots, q\}$, y que el conjunto índice de características de un ítem $I_C = \{1, 2, \dots, m\}$. En adelante denotaremos a $\text{Sim}(N(u), Ds(i))$ como Sim_i^u .

El algoritmo de recomendación consta de los siguientes pasos:

Paso 0: Filtrado PFCBI. Se aplica el proceso de pre-filtrado descrito en la sección 6.2 a los ítems cercanos (I), obteniéndose un conjunto reducido de ítems (I_R) que serán la entrada al siguiente paso del algoritmo de recomendación.

Paso 1: Para todo ítem $i \in I_R$ pre-filtrado en la fase anterior se calcula el valor Simc_i^u .

Paso 2: Suponiendo que el cardinal de I_R es n , se construye una matriz contextual, formada por los valores de similitud contextual de los ítems de I_R y las r características de contexto.

$$\text{CONTEXT}_u = \begin{pmatrix} \text{Simc}_1^u \\ \vdots \\ \text{Simc}_n^u \end{pmatrix} = \begin{pmatrix} \text{simc}_{t1}^{1-u} & \dots & \text{simc}_{tj}^{1-u} & \dots & \text{simc}_{tr}^{1-u} \\ \vdots & & \vdots & & \vdots \\ \text{simc}_{t1}^{n-u} & \dots & \text{simc}_{tj}^{n-u} & \dots & \text{simc}_{tr}^{n-u} \end{pmatrix}$$

Paso 3: Filtrado FBCC (Filtrado basado en características contextuales). Aquellos ítems que incumplan los parámetros de incompatibilidad contextual son eliminados de la lista de puntos que pueden ser recomendados. La matriz CONTEXT_u se utiliza para detectar las incompatibilidades contextuales, de tal forma que un ítem $i \in I_R$ pasará a la siguiente etapa de recomendación si cumple la siguiente ecuación:

$$\min(simc_{t1}^{i-u}, simc_{t2}^{i-u}, \dots, simc_{tr}^{i-u}) > 0$$

A este nuevo subconjunto de ítems filtrados le denominaremos I_F , que cumple que $I_F \subseteq I_R \subseteq I$.

Paso 4: Para todo ítem $i \in I_F$ debe calcularse el valor Sim_i^u . Es importante recordar que en este valor no participan los valores contextuales, solo los descriptores turísticos-demográficos y características no contextuales.

Paso 5: Se calcula para cada ítem su valor de predicción de interés para el usuario⁹⁵. Este valor de predicción se calcula de la siguiente forma:

$$pred(i, u) = \sum_{j=1}^q sim_{d_j}^{i-u} + \sum_{m=1}^c sim_{c_m}^{i-u}$$

con $j \in I_d$ y $m \in I_c$. En adelante denotaremos este valor de predicción como $pred_u^i$.

Paso 6: Ordenar la lista de ítems teniendo en cuenta el valor $pred_u^i$. En el caso de obtener para cada ítem valores diferentes entre sí, ya habría finalizado el proceso de recomendación. Si existen ítem con idénticos valores, se debe realizar un proceso de desempate que exponemos en los dos siguientes pasos.

Paso 7: Construir una matriz de interés para aquellos ítems que tienen un mismo valor $pred_u^i$ para poder desempatar y reordenarlos. Para ello

⁹⁵Se han definido en los pasos del algoritmo algunos conceptos adicionales. Se ha considerado adecuado hacerlo así ya mejora su comprensión al contextualizarla en la propia descripción del paso.

se calcula el valor $\text{pref}(i_a, i_b, u)$ que mide en cuántas características prefiere el usuario el ítem i_a al ítem i_b . En adelante lo denotaremos como $\text{pref}_{i_a i_b}$, y la forma de obtener estos valores es:

$$\text{pref}(i_a, i_b, u) = \text{pref}_{i_a i_b} = \sum_{j=1}^q \max(\text{sim}_{d_j}^{i_a-u} - \text{sim}_{d_j}^{i_b-u}, 0) + \sum_{m=1}^c \max(\text{sim}_{c_m}^{i_a-u} - \text{sim}_{c_m}^{i_b-u}, 0)$$

con $j \in I_{\mathcal{D}}$ y $m \in I_{\mathcal{C}}$

Si realizamos esos cálculos para todos los ítems empatados (n) con el mismo valor de predicción que pueden recomendarse, se obtiene una matriz de interés que nos mide las preferencias entre las diferentes alternativas por parte del usuario u . A diferencia de otros métodos (Wang, 2005; Martínet et al., 2008; Pérez, 2008), en esta matriz no intervienen todos los ítems, con lo cual el número de cálculos será bastante menor.

$$\text{PREF}_u = \begin{pmatrix} \text{pref}_{11} & \dots & \text{pref}_{1j} & \dots & \text{pref}_{1n} \\ \vdots & & \vdots & & \vdots \\ \text{pref}_{n1} & \dots & \text{pref}_{nj} & \dots & \text{pref}_{nn} \end{pmatrix}$$

Una vez calculada la matriz, es necesario ordenar los diferentes ítems según el grado de satisfacción del turista. Para ello se calcula el nivel de preferencia estricto para cada ítem. La matriz resultante la denominamos $\text{PREF}_u^{\text{estricta}}$, donde cada elemento se calcula de la siguiente manera (Wang, 2005; Martínez et al., 2008):

$$\text{pref}_{ab}^{\text{estricta}} = \max(\text{pref}_{ab} - \text{pref}_{ba}, 0)$$

Una vez obtenida esta matriz se obtiene para cada ítem un valor de grado de desempate (GD), obtenido de la siguiente forma:

$$GD_i = \sum_{j=1}^n pref_{ij} + pref_{ij}^{estricta}$$

Paso 8: Reordenar la lista de ítems. Para aquellos ítems que todavía tengan la misma puntuación se multiplica su valor de predicción $pred_u^i$ por el valor r_i^c (definición 5.9) que tenga un grado de pertenencia k (definición 5.6) mayor de entre las diferentes categorías en la que está catalogado el producto i. En caso de que varios ítems tengan el mismo valor k para varias categorías, se seleccionará el que tenga un valor r_i^c mayor. Si aún así, existieran elementos con idéntico valor se informaría que tienen el mismo grado de interés.

Otra opción a tener en cuenta en caso de tener varios ítems con el mismo valor de predicción es la aplicación de otro motor de recomendación (siguiendo un modelo de hibridación en cascada).

6.3.6. Ejemplo de aplicación del modelo propuesto

En la presente sección se detalla un ejemplo de aplicación de este motor de recomendación. Supongamos que el sistema obtiene diez ítems como resultado de aplicar el proceso PFCBI (Paso 0). Para que el ejemplo mostrado sea fácil de comprender se ha reducido el número de puntos turísticos a diez, al igual que el número de características, descriptores y atributos contextuales, aunque naturalmente en los ejemplos reales se emplean más características y mayor número de escalas. Supondremos que cada ítem viene descrito por tres atributos contextuales, tres descriptores turísticos y cinco características.

Los atributos contextuales, vendrán determinados por (t_1, t_2, t_3), los descriptores turísticos por (d_1, d_2, d_3) y las características de los puntos por (c_1, c_2, c_3, c_4, c_5). Cada una de estas características está valorada por los expertos mediante diferentes escalas lingüísticas.

En el ejemplo propuesto se han utilizado 4 escalas lingüísticas, cuyas funciones de pertenencia y sus representaciones en forma de tuplas se muestran en las siguientes figuras (Figuras 6.13 y 6.14):

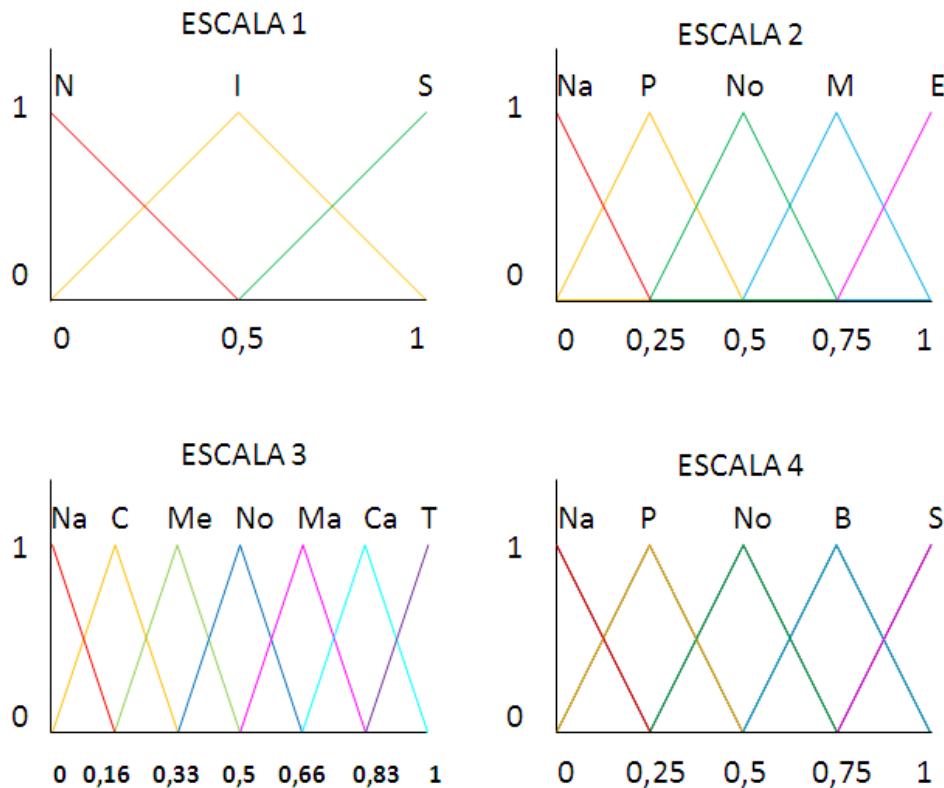


Figura 6.13. Escalas lingüísticas utilizadas en el ejemplo

ESCALA 1		ESCALA 2	
VALOR	CÓDIGO	VALOR	CÓDIGO
Sí	S (0.5, 1, 1)	Excelente	E (0.75, 1, 1)
Indiferente	I (0, 0.5, 1)	Mucho	M (0.5, 0.75, 1)
No	N (0, 0, 0.5)	Normal	No(0.25,0.5,0.75)
		Poco	P (0, 0.25, 0.5)
		Nada	Na (0, 0, 0.25)

ESCALA 3		ESCALA 4	
VALOR	CÓDIGO	VALOR	CÓDIGO
Totalmente	T (0.83, 1, 1)	Sí	S (0.75, 1, 1)
Casi todo	Ca (0.66, 0.83, 1)	Bastante	B (0.5, 0.75, 1)
Más de lo normal	Ma (0.5, 0.66, 0.83)	Normal	No(0.25,0.5,0.75)
Normal	No(0.33,0.5,0.66)	Poco	P (0, 0.25, 0.5)
Menos de lo normal	Me(0.16,0.33,0.5)	Nada	Na (0, 0, 0.25)
Casi nada	C(0, 0.16, 0.33)		
Nada	Na (0, 0, 0.16)		

Figura 6.14. Significado semántico de las escalas utilizadas en el ejemplo⁹⁶

Las características que describen los puntos turísticos son los siguientes:

	t_1	t_2	t_3	d_1	d_2	d_3	c_1	c_2	c_3	c_4	c_5
i_1	I	I	I	S	S	S	Me	N	T	No	No
i_2	I	I	I	B	B	Na	C	No	Ca	No	E
i_3	S	I	I	No	P	No	C	Na	Ca	No	No
i_4	I	I	S	P	No	No	No	P	Ma	Na	No
i_5	S	N	S	Na	P	Na	Ma	No	No	C	P
i_6	I	I	I	No	S	No	T	P	No	Me	No
i_7	N	S	N	B	S	No	Me	Na	Ca	No	P
i_8	I	I	S	Na	No	B	T	Na	T	T	Na
i_9	N	I	S	P	No	B	Me	E	Ca	No	E
i_{10}	I	I	N	No	B	B	Na	P	No	Ma	P

Tabla 6.4. Descripción de puntos turísticos

⁹⁶ Las funciones de pertenencia se representan mediante una 3-tupla (a,b,c) donde b es el valor donde la función de pertenencia tiene valor 1, mientras que a y c indican los extremos izquierdo y derecho de la función

Concretamente, los atributos contextuales han utilizado la escala número 1 y para los descriptores, la escala 4. Por último, se ha empleado la escala 3 para las características primera, tercera y cuarta y la escala 2 para describir la segunda y la quinta.

El sistema requiere al turista expresar sus preferencias y necesidades respecto a los puntos turísticos que desea que le sean recomendados (Tabla 6.5).

	t_1	t_2	t_3	d_1	d_2	d_3	c_1	c_2	c_3	c_4	c_5
u_1	S	N	I	No	B	B	Na	P	No	No	Na

Tabla 6.5. Descripción de necesidades del usuario

Paso 1. Se calcula el vector de similitud entre las necesidades del usuario y cada uno de los ítems.

$$Simc(N(u), Ds(i))\} = Simc_i^u$$

$$Simc_1^u = (0.5, 0.5, 1) \quad Simc_2^u = (0.5, 0.5, 1)$$

$$Simc_3^u = (1, 0.5, 1) \quad Simc_4^u = (0.5, 0.5, 0.5)$$

$$Simc_5^u = (1, 1, 0.5) \quad Simc_6^u = (0.5, 0.5, 1)$$

$$Simc_7^u = (0, 0, 0.5) \quad Simc_8^u = (0.5, 0.5, 0.5)$$

$$Simc_9^u = (0, 0.5, 0.5) \quad Simc_{10}^u = (0.5, 0.5, 0.5)$$

Paso 2. Construcción de la matriz contextual.

$$CONTEXT_u = \begin{pmatrix} 0.5 & 0.5 & 1 \\ 0.5 & 0.5 & 1 \\ 1 & 0.5 & 1 \\ 0.5 & 0.5 & 0.5 \\ 1 & 1 & 0.5 \\ 0.5 & 0.5 & 1 \\ 0 & 0 & 0.5 \\ 0.5 & 0.5 & 0.5 \\ 0 & 0.5 & 0.5 \\ 0.5 & 0.5 & 0.5 \end{pmatrix}$$

Paso 3. Filtrado FBCC.

Los puntos que incumplen la incompatibilidad contextual son los puntos turísticos i_7 e i_9 , ya que

$$\min(simc_{t1}^{7-u}, simc_{t2}^{7-u}, simc_{t3}^{7-u}) = 0 \text{ y } \min(simc_{t1}^{9-u}, simc_{t2}^{9-u}, simc_{t3}^{9-u}) = 0$$

Paso 4. Para los ocho ítems filtrados se calcula el valor $Sim(N(u), Ds(i)) = Sim_i^u$

Así, los valores obtenidos serían:

$$Sim_{i1}^u = (0, 0.5, 0.5 | 0, 0.5, 0, 1, 0) \quad Sim_{i2}^u = (0.5, 1, 0 | 0.5, 0.5, 0, 1, 0)$$

$$Sim_{i3}^u = (1, 0, 0.5 | 0.5, 0.5, 0, 0.5, 0) \quad Sim_{i4}^u = (0.5, 0.5, 0.5 | 0, 1, 0.5, 0, 0)$$

$$Sim_{i5}^u = (0, 0, 0 | 0, 0.5, 1, 0, 0.5) \quad Sim_{i6}^u = (1, 0.5, 0.5 | 0, 1, 1, 0.5, 0)$$

$$Sim_{i8}^u = (0, 0.5, 1 | 0, 0.5, 0, 0, 1) \quad Sim_{i10}^u = (1, 1, 1 | 1, 1, 1, 0.5, 0.5)$$

Paso 5: Para los ocho ítems se calcula su valor de predicción de interés para el usuario.

$$pred(i, u) = \sum_{j=1}^q sim_{d_j}^{i-u} + \sum_{m=1}^c sim_{c_m}^{i-u}$$

$$pred(i_1, u) = 1 + 1.5 = 2.5$$

$$pred(i_2, u) = 1.5 + 2 = 3.5$$

$$pred(i_3, u) = 1.5 + 1.5 = 3 \quad pred(i_6, u) = 2 + 2.5 = 4.5$$

$$pred(i_4, u) = 1.5 + 1.5 = 3 \quad pred(i_8, u) = 1.5 + 1.5 = 3$$

$$pred(i_5, u) = 0 + 2 = 2 \quad pred(i_{10}, u) = 3 + 4 = 7$$

Paso 6: Ordenar la lista de ítems teniendo en cuenta el valor $pred_u^i$. En este caso el orden sería:

$$i_{10}, i_6, i_2, i_3, i_4, i_8, i_1, i_5$$

Como puede observarse los puntos i_2, i_3, i_4 están empataados por lo que a estos puntos se les aplican los siguientes pasos.

Paso 7. Construir una matriz de interés y de interés estricto para obtener los grados de desempate.

$$pref(a, b, u) = pref_{ab} = \sum_{j=1}^q \max(sim_{d_j}^{a-u} - sim_{d_j}^{b-u}, 0) + \sum_{m=1}^c \max(sim_{c_m}^{a-u} - sim_{c_m}^{b-u}, 0)$$

A continuación, se muestra el algoritmo aplicado para su cálculo.

```

int a.b.k;
float gradopreferencia, temporal;
Para a desde 1 hasta numpuntosturisticosempatados con incremento 1
Para b desde 1 hasta numpuntosturisticosempatados con incremento 1
{
    gradopreferencia=0;
    Para k desde 1 hasta ATR con incremento 1
    //ATR es la suma del número de descriptores y características de los puntos
    {
        temporal=sim(k, a, u) - sim(k, b, u) ]
        if (temporal<0)
            temporal=0;
        gradopreferencia+=temporal;
    }
    pref(a, b, u) =gradopreferencia;
}

```

Así, la matriz de preferencia sería:

$$PREF_u = \begin{pmatrix} 0 & 1.5 & 2 \\ 1 & 0 & 1.5 \\ 1.5 & 1.5 & 0 \end{pmatrix}$$

Partiendo de la matriz de preferencia, se calcula el nivel de preferencia estricto de cada ítem. Para ello se aplica el proceso siguiente algoritmo:

```

int a,b;
float temporal;
Para a desde 1 hasta numpuntosturisticosempatados con incremento 1
Para b desde 1 hasta numpuntosturisticosempatados con incremento 1
{
    temporal=pref(a,b,u)-pref(b,a,u)
    si (temporal>0)
        prefestricta(a,b,u) =temporal;
    en caso contrario
        prefestricta(a,b,u) = 0;
}

```

La matriz de preferencias estrictas obtenida sería:

$$PREF_u^{estricta} = \begin{pmatrix} 0 & 0.5 & 0.5 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Una vez obtenida esta matriz obtenemos para cada ítem el valor de grado de desempate (GD), obteniendo como resultados los siguientes valores:

i_2	i_3	i_4
4.0	2.5	3.0

Tabla 6.6. Grado de desempate de los ítems

Paso 8. Se reordena la lista de ítems, quedando los cinco primeros:

$$i_{10}, i_6, i_2, i_4, i_3$$

Debido a que no quedan ítems con el mismo grado de desempate , no se haría necesario tener que utilizar los valores r_i^c

6.3.7. Consideraciones finales

Dadas las características y la forma en que se ha definido el motor de recomendación presentado en este apartado, se puede afirmar que su aplicación es adecuada en las siguientes situaciones:

- a) Cuando el usuario expresa sus preferencias explícitamente.
- b) Cuando el usuario expresa sólo alguna de las características de los ítems que desea visitar, tomando el resto de características el valor indiferente en la escala lingüística correspondiente.
- c) Cuando el usuario expresa sus preferencias seleccionando como referencia un ítem visitado anteriormente. En este caso diríamos estaríamos hablando de un motor de recomendación basado en contenido con memoria⁹⁷.
- d) Cuando varios usuarios realizan visitas en grupo⁹⁸ y las preferencias de cada uno de los miembros del grupo se obtienen de alguna de las tres maneras indicadas en los apartados a, b y c.

⁹⁷ El motor de recomendación basado en contenido con memoria sería exactamente igual al propuesto en esta sección, solo que la información respecto a las necesidades del usuario provienen de una visita realizada anteriormente.

⁹⁸ En la sección 6.7. se detalla la forma de realizar la recomendación grupal.

El motor propuesto resulta de especial interés en aquellos casos en los que no se dispone de información de visitas anteriores⁹⁹. El utilizar información difusa resulta muy interesante cuando el turista no especifica exactamente sus preferencias, haciendo que las no descritas tomen un valor indiferente. En el caso de requerirse un método de recomendación más exhaustivo podrían calcularse las matrices de preferencias con todos los ítems (empatados y no empatados), seleccionando el ítem de mayor valor GD. Posteriormente, se repetiría este mismo paso, hasta que se tengan el número buscado de ítems, aunque eliminando el ítem de mayor grado. En ese momento el proceso finalizaría.

Esta solución tiene el inconveniente de que requiere un número mayor de cálculos y los resultados obtenidos en diferentes pruebas realizadas son idénticos en un 95%. Otra posible mejora podría ser que el valor r_i^c , que se utiliza en el paso 8, tuviera en cuenta las evaluaciones, según el contexto aplicado o los valores promedios de las diferentes valoraciones en las que intervengan los atributos contextuales, ya que estos valores están almacenados en el sistema.

6.4. Motor de recomendación basado en técnicas demográficas (SGP^u)

Los motores de recomendación demográficos, como se describió en el capítulo tercero, realizan las recomendaciones teniendo en cuenta el grupo demográfico o estereotipo al que pertenece el usuario (Vozalis y Margaritis, 2007; Pérez, 2008). Por tanto, antes de realizar el proceso de recomendación el usuario debe estar catalogado en alguno de los grupos

⁹⁹ Problema muy habitual en muchos sistemas de recomendación, tal como se describió en el capítulo 3.

predefinidos, para tener sólo en cuenta las valoraciones realizadas por aquellos usuarios categorizados de manera idéntica o muy similar.

Por tanto, para un correcto funcionamiento, será necesario disponer de las valoraciones realizadas por otros turistas. Así, cuando se guarda una valoración sobre un ítem, no solo debe almacenarse esta valoración sino las características turístico-demográficas de la persona que ha realizado dicha valoración.

En el sistema que se propone en este trabajo se incluye un motor de recomendación demográfico. Éste aplica un proceso consistente en detectar los diferentes ítems o puntos turísticos que se encuentran a una distancia menor a una dada del punto en el que está localizado el turista, para posteriormente calcular las valoraciones medias obtenidas de cada ítem, partiendo de las valoraciones otorgadas por los turistas pertenecientes al mismo estereotipo o similares¹⁰⁰. Una de las características de este motor es que no es necesario que el usuario solicite ningún tipo de característica acerca del punto de interés o actividad que desea visitar.

Un hecho frecuente es que las valoraciones de un mismo ítem o actividad turística, pueden ser bastante diferentes entre distintos turistas de un mismo grupo. Por este motivo, se toma como referencia el valor medio.

Cada turista es descrito mediante una serie de descriptores o atributos. Dado que suele ser complejo medir ciertas características con

¹⁰⁰ En las definiciones 6.14 y 6.15 explicamos cómo mediremos la similaridad entre dos estereotipos.

un valor cuantitativo, consideramos que lo mejor en estos casos es utilizar una descripción cualitativa que posteriormente se traslada a una escala numérica de forma automática. En nuestra propuesta, en el proceso de registro el turista debe describir sus características mediante escalas lingüísticas adecuadas. Puesto que las distintas características pueden ser muy heterogéneas, es adecuado el uso de diferentes escalas.

La mayoría de los sistemas demográficos generan recomendaciones de puntos turísticos muy similares, por lo que uno de los inconvenientes que presentan es que no es fácil que se produzca el efecto sorpresa (Fesenmaier et al, 2006). En el modelo que se propone se pretende que este motor sea capaz de aportar un efecto sorpresa al usuario sin perder la esencia del funcionamiento habitual de un sistema de recomendación demográfico.

6.4.1. Descripción.

Como paso previo a la ejecución del motor de recomendación propuesto en esta sección, se debe producir la fase de pre-filtrado contextual basado en implicaciones (PFCBI¹⁰¹). De este modo, partiendo de información contextual del usuario y mediante implicaciones se reduce de forma muy significativa el número de ítems con los que se trabaja, reduciendo el tiempo de ejecución requerido por el motor de recomendación.

Posteriormente, se realiza una comparación de los atributos contextuales del turista respecto a los de los diferentes ítems o puntos

¹⁰¹ Sección 6.2.

turísticos que pueden recomendarse. De esta manera, se excluyen aquellos puntos que no sean compatibles con el contexto actual, manteniendo únicamente aquellos ítems que tienen una alta compatibilidad con el contexto actual del usuario que solicita la recomendación.

La mayoría de los sistemas de recomendación demográficos se basan exclusivamente en las valoraciones de turistas que pertenecen a la misma categoría del usuario que solicita la recomendación, mientras que nuestro modelo tendrá en cuenta además otras consideraciones.

Una vez realizada las etapas de pre-filtrado descritas anteriormente, se utilizarán dos medidas para cada ítem i:

- a) El valor medio de confianza¹⁰² (*denotado* v_{i-t}^{conf}), que es la valoración media realizada por turistas que pertenecen al mismo estereotipo o a estereotipos similares o vecinos¹⁰³ al turista t.
- b) La valoración general¹⁰⁴ (*denotada* C_i) realizada por todos los turistas (pertenezcan o no al estereotipo del turista que solicita la recomendación) del ítem i. Se trata de un valor basado en las valoraciones medias que realizan los turistas que visitan dicho ítem.

Por tanto, el modelo propuesto utiliza para su funcionamiento no sólo valoraciones de turistas que pertenecen a un mismo grupo demográfico o similar, sino de todos los turistas que visitan un ítem que

¹⁰² Definición 6.19

¹⁰³ Definición 6.16 y 6.17

¹⁰⁴ Definición 6.20

puede ser recomendado. Este aspecto nos resultará útil para aportar al modelo el efecto sorpresa que pretendemos, ya que cuando un turista realiza una valoración de una actividad turística, independientemente del estereotipo al que pertenezca, estará influyendo en las futuras recomendaciones que el motor generará.

Partiendo del perfil demográfico del turista, se obtiene información sobre sus características más importantes. Dado que se quiere provocar un efecto sorpresa, de cada ítem vamos a considerar una valoración determinada que es la que servirá de referencia a la hora de calcular el valor de predicción de interés que tendrá para el turista que solicita la recomendación dicho ítem. En vez de considerar solo las valoraciones realizadas por turistas que pertenezcan al mismo estereotipo, en nuestro modelo no haremos aquí ninguna discriminación.

Así, para cada ítem se tomará en consideración la evaluación cuyo valor sea el más pequeño de todas las que superen la valoración media de confianza (v_{i-t}^{conf}).

El criterio de selección de la visita que se va a tomar como referencia para poder generar los valores de predicción para la recomendación, podría entenderse como un atributo configurable. Así, podría elegirse por ejemplo, la mejor valoración del ítem, o bien seleccionar varias valoraciones con una calificación media igual a ese mismo valor mínimo, o incluso varias evaluaciones de diferentes turistas aunque tengan valoraciones diferentes.

En nuestro caso, para simplificar el modelo, se ha considerado utilizar solamente una valoración por cada ítem. Como ya se ha indicado, concretamente se selecciona aquella que tenga el valor mínimo de entre las que superen la valoración media de confianza.

Dado que las valoraciones seleccionadas pueden corresponder a turistas que pertenezcan a estereotipos diferentes al turista que solicita la recomendación, los pasos posteriores van encaminados a calcular cuál de los turistas que han realizado esas valoraciones seleccionadas tiene un perfil más similar al turista que solicita la recomendación. Es decir, las recomendaciones que generará el motor cumplirán que tienen una buena calificación y que los turistas que realizaron la valoración tienen un grado de afinidad con el turista que solicita la recomendación, pero no tienen forzosamente que pertenecer al mismo estereotipo.

Uno de los inconvenientes de la mayoría de los motores de recomendación demográficos es que en muchas ocasiones puede ser recomendado un ítem que está peor valorado que otros, pero que son visitados por turistas que tienen un perfil igual o similar al que solicita la recomendación.

En cambio el modelo que se propone en este trabajo recomendará ítems con alta valoración, pero que quizás no son normalmente visitados por usuarios con perfil idéntico al que solicita la recomendación. Esta es la manera de intentar provocar un efecto sorpresa en las recomendaciones, es decir, hacer que en la lista de ítems recomendados aparezca algún ítem “inesperado” para el turista que solicita la recomendación.

El mecanismo que aplica nuestro modelo para eliminar ítems con baja valoración, consiste en eliminar todos aquellos que tengan una calificación general inferior a la media de las valoraciones. De esta forma, la recomendación va a realizarse sólo sobre ítems que han sido bien valorados.

Una vez excluidos aquellos productos que tengan una valoración inferior a la valoración media de las diferentes actividades que pueden ser recomendadas, se realiza un proceso en el que se calcula la similitud entre el perfil del turista que ha realizado la valoración seleccionada y el perfil del turista que solicita la recomendación. Una vez obtenida esta información, el sistema recomendará aquellos puntos que hayan obtenido un grado de similitud mayor entre el turista que solicita la recomendación y el que realizó la valoración, teniendo en cuenta también la valoración general del ítem C_i y la valoración de confianza v_{i-t}^{conf} .

6.4.2. Formalización

Sea I el conjunto de ítems situados a una distancia menor o igual a la indicada por el turista.

$$I = \{i_1, i_2, i_3, \dots, i_n\}$$

Sea U el conjunto de usuarios que utilizan el sistema propuesto.

$$U = \{u_1, u_2, u_3, \dots, u_m\}$$

Sea $I_{\mathcal{D}}$ el conjunto índice de descriptores turísticos y demográficos {1, 2,d}. A cada dominio D_j con $j \in I_{\mathcal{D}}$ le añadiremos un nuevo elemento

denominado ‘indiferente’ y denotado \perp que representará que no se dispone de información acerca del descriptor d_j .

Definición 6.12: Se denomina perfil turístico-demográfico del turista u , denotado $P(u)$, como una tupla de valores de descriptores turísticos-demográficos (d) que lo define. Más formalmente:

$$P: U \rightarrow \prod_{j \in D} D_j$$

Por tanto, el perfil $P(u)$ sería:

$$P(u) = (d_1^u, d_2^u, \dots, d_q^u)$$

donde $d_j^u (j \in I_{\mathcal{D}})$, es el valor del descriptor j para el turista u . Esta información es proporcionada por el turista.

Definición 6.13: Cada ítem o actividad a recomendar i_k , tiene un conjunto de valoraciones, donde cada una de ellas tiene asociada un valor numérico y una lista de características o descriptores turísticos que describen al turista t que lo ha valorado. A esta información se le llama valoración general de la actividad i_k del turista t y se formaliza como $V(i_k, t)$ y denotado como $V_{i_k}^t$:

$$V: IxU \rightarrow \mathbb{R} \times \prod_{j \in D} D_j$$

$$V(i_k, t) = V_{i_k}^t = (v_{i_k}^t, d_1^t, d_2^t, \dots, d_q^t)$$

donde d_j^t (con $j \in I_{\mathcal{D}}$), es el valor del descriptor j para el turista t y $v_{i_k}^t$ representa la valoración general que ha tenido por parte del turista t la visita o realización de la actividad i_k .

Definición 6.14: Se denomina estereotipo (denotado E) al conjunto de todas las combinaciones de los descriptores turísticos-demográficos (dp).

Por ejemplo, si suponemos la existencia de tres descriptores turísticos d_1 , d_2 y d_3 , y suponemos que el primero tiene tres posibles valores (0, 1 y 2), el segundo cuatro posibles valores (0, 1, 2 y 3) y el tercero dos valores (0 y 1), el número total de estereotipos que podrían asignarse sería 24. En la tabla 6.7 podemos observar todos los estereotipos existentes:

	d_1	d_2	d_3		d_1	d_2	d_3
E_1	0	0	0	E_{13}	1	2	0
E_2	0	0	1	E_{14}	1	2	1
E_3	0	1	0	E_{15}	1	3	0
E_4	0	1	1	E_{16}	1	3	1
E_5	0	2	0	E_{17}	2	0	0
E_6	0	2	1	E_{18}	2	0	1
E_7	0	3	0	E_{19}	2	1	0
E_8	0	3	1	E_{20}	2	1	1
E_9	1	0	0	E_{21}	2	2	0
E_{10}	1	0	1	E_{22}	2	2	1
E_{11}	1	1	0	E_{23}	2	3	0
E_{12}	1	1	1	E_{24}	2	3	1

Tabla 6.7. Ejemplo de estereotipos

Definición 6.15: Dado $j \in I_{\mathcal{D}}$, se denomina escala del descriptor j y se denota como E^j , al conjunto de posibles valores o términos lingüísticos definidos para valorar el descriptor j.

$$E^j = \{e_1^j, e_2^j, \dots, e_z^j\}$$

Por tanto, el valor del descriptor j por parte del turista u debe cumplir:

$$d_j^u \in E^j, \text{ para todo } j \in I_{\mathcal{D}}$$

Para cada descriptor turístico supondremos que los términos forman una estructura ordenada que cumple:

$$e_i < e_j \Leftrightarrow i < j \text{ y } Neg(e_i) = e_j, \text{ con } j = z - i + 1$$

Para poder realizar operaciones que sirvan para determinar si un turista puede ser considerado de un estereotipo similar al que pertenece se necesita definir una función de transformación que asigne a cada valor lingüístico un valor numérico igual a la posición que ocupa el término lingüístico. Sea A un conjunto de términos lingüísticos que describe un descriptor turístico:

$$A = \{e_1, e_2, \dots, e_z\}$$

Se define la función de transformación:

$$F_{TRANS}: A \rightarrow \mathbb{N} / F_{TRANS}(e_i) = i$$

donde i es la posición que ocupa el término lingüístico. Se denomina rango de un descriptor j (denotado como $rango(d_j)$) al número de términos lingüísticos que tiene el descriptor j, es decir, la cardinalidad.

Definición 6.16: Dos usuarios u_1 y u_2 pertenecen al mismo estereotipo o grupo si todos sus descriptores turísticos-demográficos son idénticos.

Definición 6.17: Dos usuarios u_1 y u_2 pertenecientes a distintos estereotipos son vecinos o tienen estereotipos similares, si cumplen la siguiente condición:

$$P(u1) = (d_1^{u1}, d_2^{u1}, \dots, d_n^{u1})$$

$$P(u2) = (d_1^{u2}, d_2^{u2}, \dots, d_n^{u2})$$

$$\sum_{i=1}^n \frac{|F_{TRANS}(d_i^{u1}) - F_{TRANS}(d_i^{u2})|}{rango(d_i)} < umbral$$

En un ejemplo posterior ilustraremos esta definición usando como umbral el valor (sección 6.4.4):

$$\sum_{i=1}^n \frac{1}{rango(d_i)}$$

6.4.3. Proceso de filtrado y recomendación

Para cada ítem del conjunto generado en la fase PFCBI debe realizarse un proceso de filtrado y recomendación, teniendo en cuenta el estereotipo al que pertenecen los turistas que han visitado los puntos turísticos y sus valoraciones. Para ello, se calcula la valoración media de todas los ítems, descartando aquellos con un valor inferior al valor medio obtenido.

Definición 6.18: Denominamos valoración de confianza del ítem i_k para el turista t , (denotada $v_{i_k-t}^{conf}$) a la media de las valoraciones de i_k

correspondientes a turistas que pertenecen al mismo estereotipo o a estereotipos vecinos del turista t que solicita la recomendación.

Definición 6.19: Se define la calificación media del ítem i_k , denotada C_{i_k} a la valoración media dada por los turistas que han visitado la actividad i_k y que cumplan el criterio que a continuación se describe.

Para la obtención de esta calificación media se eliminan el 15% de la calificaciones más altas y el 15% de las calificaciones más bajas. De esta forma, se intenta evitar sesgos generados por usuarios que al calificar pretendan alterar la calificación otorgada. Una vez eliminadas estas valoraciones extremas, sólo se utiliza para calcular C_{i_k} el 70% restante de las visitas valoradas. Esta información debe estar almacenada en las bases de datos del sistema con objeto de no tener que calcularla cada vez que se realice una petición de recomendación.

Definición 6.20: Se define C_{avg} como la media aritmética de las valoraciones media C_i de las n actividades i pertenecientes a I prefiltradas en la fase PFCBI. Su valor se obtiene de la siguiente manera:

$$C_{avg} = \frac{1}{n} \sum_{r=1}^n C_{i_r}$$

Como ya se ha indicado, sólo participarán en el proceso de recomendación aquellas actividades i_k que cumplan la siguiente regla:

$$C_{i_k} \geq C_{avg}$$

Definición 6.21: La valoración de la actividad i_k seleccionada para realizar el proceso de recomendación (denotada $V_{i_k}^{select}$) es la valoración de i_k realizada por un turista cuyo valor es el mínimo que supera a C_{i_k}

(es decir, si todas las valoraciones están ordenadas de forma ascendente sería la siguiente que supera a C_{i_k}), siendo C_{i_k} la calificación media de todas las valoraciones otorgadas por los turistas que han visitado la actividad i_k , y que se han considerado válidas para su cálculo.

Esta información debe tenerla el sistema calculada para cada ítem. Un caso especial se da cuando coinciden varios ítems con el mismo valor de $V_{i_k}^{select}$. En este caso participará en el proceso de recomendación la valoración de la actividad turística o ítem i_k dada por el turista que se asemeje más al que solicita la recomendación. Es decir, si por ejemplo tenemos un ítem cuyo C_{i_k} tiene un valor de 7.6 y existen dos visitantes con una calificación 7,7 ¿Cuál de las dos visitas se utilizará en el proceso de recomendación? En nuestro modelo aquella que pertenezca a un turista que tenga un perfil más similar al turista que solicita la recomendación.

Sea I_γ el conjunto índice de características contextuales {1, 2, ...t}, a cada dominio D_i con $i \in I_\gamma$ se le añade un nuevo elemento denominado ‘indiferente’ y denotado \perp que representará que no se dispone de información acerca de la característica contextual.

Definición 6.22: Se define la necesidad contextual del turista u , denotado $Nct(u)$, como una tupla formada por los valores de atributos contextuales del turista. Más formalmente:

$$Nct: U \rightarrow \prod_{i \in T} D_i$$

Por tanto

$$Nct(u) = (t_1^u, t_2^u, \dots, t_t^u)$$

Donde t_i^u , es el valor del atributo contextual i para el usuario u con $i \in I_{\tau}$. Esta información es proporcionada por el turista.

Definición 6.23. Se define el concepto de descriptor contextual del ítem i_k , denotado, $Dsct(i_k)$, como una tupla formada por sus valores de atributos contextuales (t). Más formalmente:

$$Dsct: I \rightarrow \prod_{i \in T} D_i$$

El descriptor $Dsct(i_k)$ es por tanto

$$Dsct(i_k) = (t_1^{ik}, t_2^{ik}, \dots, t_t^{ik})$$

Donde t_i^{ik} es el valor del atributo contextual i para la actividad o ítem i_k , con $i \in I_{\tau}$. Todos estos valores han de ser proporcionados por los expertos del destino.

Definición 6.24: Se define la similitud contextual en el atributo k del ítem i y del contexto del usuario u ($simc(k,i,u)$), a la medida de la similitud de los valores lingüísticos del atributo contextual k asociados al ítem i y al usuario u . Al valor $simc(k,i,u)$, lo denominaremos alternativamente como $simc_k^{i-u}$. Para calcular este valor se aplicará la medida de posibilidad en la que dadas dos funciones de pertenencia A y B se define como:

$$Poss(A,B) = \sup_{x \in X} \{ \min(A(x), B(x)) \}$$

Definición 6.25: Se denomina similitud contextual del ítem i y del usuario u $Simc(Nct(u), Dsct(i))$, denotada como $Simc_i^u$, a la lista de los

valores de similitud para todos los atributos contextuales lingüísticos.

Formalmente,

$$Simc(Nct(u), Dsct(i)) = (simc_{t1}^{i-u}, simc_{t2}^{i-u}, \dots, simc_{tr}^{i-u})$$

donde r pertenece a I_{τ}

Definición 6.26: Se define la similitud del descriptor lingüístico k de un ítem i para dos turistas t_1 y t_2 , denotado $sim(k, i, t_1, t_2)$ como el valor que mide la similitud entre ambos valores lingüísticos, utilizando para ello la medida de posibilidad.

En nuestro proceso de recomendación, este valor lo calcularemos para el turista t que ha evaluado el ítem o actividad turística i y el usuario u que solicita la recomendación. Al valor $sim(k, i, t, u)$ lo denotaremos en adelante como $sim_k^i(t-u)$.

Definición 6.27: Se define la similitud del turista u (que ha solicitado una recomendación) con respecto al turista t (que ha valorado el punto turístico i), como la lista obtenida a partir de la comparación de todos los descriptores lingüísticos.

$$Sim(P(t), P(u), i) = (sim_{d1}^{i(t-u)}, sim_{d2}^{i(t-u)}, \dots, sim_{dq}^{i(t-u)})$$

donde I_d es el conjunto índice de descriptores turísticos y demográficos $\{1, 2, \dots, q\}$. En adelante denotaremos a $Sim(P(t), P(u), i)$ como Sim_{t-i}^u .

Definición 6.28: Se denomina idoneidad de una valoración realizada por el turista t del ítem k respecto al turista u como el valor que

calcula cómo de parecido es el turista t que valoró el ítem k respecto al usuario u que solicita la recomendación. La forma de calcularlo será:

$$idon(t, k, u) = \sum_{j=1}^q sim_{d_j}^{k \rightarrow (t-u)}$$

donde q es el número de descriptores turísticos-demográficos.

Una vez realizadas estas definiciones pasamos a describir los pasos del algoritmo de recomendación:

Paso 0: Filtrado PFCBI (Pre-filtrado contextual basado en implicaciones), que (véase la Sección 6.2) a partir del conjunto de ítems situados en un rango próximo al turista (I), aplica una fase de pre-filtrado contextual, obteniéndose un conjunto reducido de ítems (I_R) que será la entrada al siguiente paso del algoritmo de recomendación.

Paso 1: Para todo ítem pre-filtrado de la fase anterior se calcula el valor $Simc_i^u$

Paso 2: Suponiendo que el cardinal de I_R es n se construye una matriz contextual, formada por los valores de similitud contextual de los ítems de I_R y las r características de contexto.

$$CONTEXT_u = \begin{pmatrix} Simc_1^u \\ \vdots \\ Simc_n^u \end{pmatrix} = \begin{pmatrix} simc_{t1}^{1-u} & \dots & simc_{tj}^{1-u} & \dots & simc_{tr}^{1-u} \\ \vdots & & \vdots & & \vdots \\ simc_{t1}^{n-u} & \dots & simc_{tj}^{n-u} & \dots & simc_{tr}^{n-u} \end{pmatrix}$$

Paso 3: Filtrado FBCC. Aquellos ítems o actividades turísticas que incumplan los parámetros contextuales son eliminados de la lista de ítems que pueden ser recomendados. La matriz $CONTEXT_u$ se utiliza para detectar las incompatibilidades contextuales, de tal forma que un

ítem $i \in I_R$, pasará a la siguiente etapa de recomendación si cumplen la siguiente ecuación:

$$\min(simc_{t1}^{i-u}, simc_{t2}^{i-u}, \dots, simc_{tr}^{i-u}) > 0$$

A este nuevo subconjunto de ítems filtrados le denominaremos I_F , que cumple que $I_F \subseteq I_R \subseteq I$.

Paso 4: Para todo ítem $i_k \in I_F$ que el sistema ha filtrado en el paso anterior y que cumple la condición $C_{i_k} \geq C_{avg}$, se obtiene el valor $V_{i_k}^{select}$.

Paso 5. Se calcula la similitud entre las características del turista t que solicita la recomendación y el turista u seleccionado para cada ítem $i_k \in I_F$. Por tanto para toda $V_{i_k}^{select}$ seleccionada en el paso anterior debe calcularse el valor $Sim_{t-i_k}^u$, donde $t-i_k$ es el turista que generó el valor $V_{i_k}^{select}$

En este punto se tiene una lista de ítems y para cada uno de ellos una valoración seleccionada generada por un turista dado. En el siguiente paso se calcula cuál de esos turistas es más parecido al turista que solicita la recomendación.

Paso 6: Obtener el valor de idoneidad de la valoración del turista t que ha generado la valoración seleccionada del ítem k , para de esta forma aplicando la definición 6.28, conocer la similitud con el turista que realiza la petición de recomendación (u).

$$idon(t, k, u) = \sum_{j=1}^q sim_{d_j}^k (t-u)$$

Paso 7: Calcular el grado de predicción de cada ítem por el usuario.

Para cada ítem i pre-filtrado en el paso 4, se calculará su valor de predicción para el usuario u (pred_u^i) de la siguiente manera:

$$\text{pred}_u^i = \text{idon}(t, i, u) \cdot V_i^{\text{select}} + v_{i-u}^{\text{conf}}$$

Un motor demográfico utilizaría como valor de predicción exclusivamente el valor v_{i-u}^{conf} . En nuestro modelo para añadir el efecto sorpresa, se añade a este valor el grado de idoneidad con la valoración general que tiene la actividad.

Paso 8: Reordenar la lista de ítems. En caso de que existan ítems con el mismo valor de predicción, se seleccionará el que tenga un valor v_{i-u}^{conf} mayor. Si aún así, existieran elementos con idéntico valor se informaría que tienen el mismo grado de interés para el turista.

6.4.4. Ejemplo de aplicación del modelo propuesto

En esta sección presentamos un ejemplo práctico de aplicación del modelo propuesto para el motor de recomendación demográfica.

Supóngase que tras la aplicación del proceso PFCBI se obtiene un conjunto de diez ítems. Para simplificar, consideraremos la existencia de sólo tres atributos contextuales y tres descriptores. Asimismo, presentaremos directamente las valoraciones, para no abundar demasiado en los cálculos. Además, ilustraremos especialmente el concepto de vecindad entre estereotipos (definición 6.17).

Consideraremos los atributos contextuales t_1 , t_2 y t_3 , y los descriptores turísticos-demográficos d_1 , d_2 y d_3 que definen el estereotipo.

En el ejemplo se han utilizado dos escalas lingüísticas que se representan en la siguiente figura junto a la representación en forma de tupla de las respectivas funciones de pertenencia (Figura 6.15):

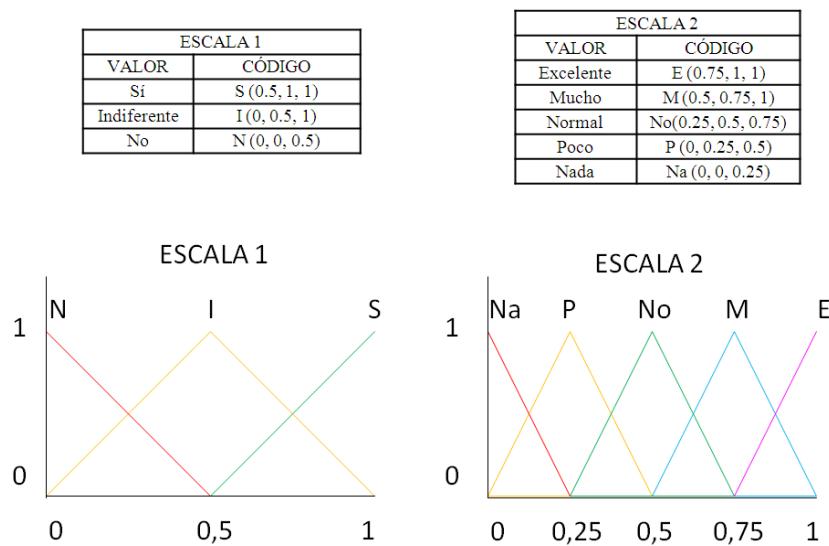


Figura 6.15. Escalas lingüísticas y significado semántico

Concretamente, para los atributos contextuales se utiliza la escala 1 y para los descriptores la escala 2. Supongamos que los usuarios u_1 , u_2 y u_3 están definidos por descriptores de la siguiente manera:

$$u_1 = (E, M, Na)$$

$$u_2 = (E, P, Na)$$

$$u_3 = (E, P, P)$$

Teniendo en cuenta que la función de transformación en esta escala tendría los siguientes valores en las diferentes etiquetas:

$$F_{TRANS}(E) = 5$$

$$F_{TRANS}(M) = 4$$

$$F_{TRANS}(No) = 3$$

$$F_{TRANS}(P) = 2$$

$$F_{TRANS}(Na) = 1$$

Los usuarios u_1 y u_2 son vecinos porque:

$$\frac{|5 - 5|}{5} + \frac{|4 - 2|}{5} + \frac{|1 - 1|}{5} = 0,2 < umbral = \frac{1}{5} + \frac{1}{5} + \frac{1}{5} = 0,6$$

Los usuarios u_1 y u_3 no son vecinos porque:

$$\frac{|5 - 5|}{5} + \frac{|4 - 2|}{5} + \frac{|1 - 1|}{5} = 0,6 \not< umbral = 0,6$$

Los usuarios u_2 y u_3 son vecinos porque:

$$\frac{|5 - 5|}{5} + \frac{|2 - 2|}{5} + \frac{|1 - 2|}{5} = 0,2 < umbral = 0,6$$

Los atributos contextuales, la calificación media (C_{i_k}) y la valoración de confianza (v_i^{conf}) que describen los puntos turísticos son los siguientes¹⁰⁵:

¹⁰⁵ Se suponen ya calculados para facilitar la comprensión del ejemplo.

	t_1	t_2	t_3	C_{i_k}	v_i^{conf}
i_1	I	I	I	7,25	6,4
i_2	I	I	I	8,48	7,6
i_3	S	I	I	6,25	8
i_4	I	I	S	6,8	5,6
i_5	S	N	S	7,8	6,1
i_6	I	I	I	7,1	7,2
i_7	N	S	N	5,8	6,3
i_8	I	I	S	8,2	8,7
i_9	N	I	S	4,3	6,7
i_{10}	I	I	N	7,9	7,7

Tabla 6.8. Descripción de puntos turísticos

Supongamos que el turista que solicita la recomendación está definido por los siguientes parámetros contextuales y descriptores turísticos.

	t_1	t_2	t_3	d_1	d_2	d_3
u	S	N	I	M	M	M

Tabla 6.9. Descripción contextual y estereotipo del usuario

Una vez presentada la información sobre los diferentes ítems y las características demográficas del turista que solicita la recomendación se muestra la aplicación del algoritmo:

Paso 0. Se aplica el proceso FPCBI para reducir el número de puntos turísticos a los que aplicar el motor de recomendación. Suponemos que

en esta fase se obtiene como resultado los diez ítems anteriormente descritos.

Paso 1. Se crea para cada ítem la lista $Simc_i^u$

$$Simc_{i1}^u = (0.5, 0.5, 1) \quad Simc_{i2}^u = (0.5, 0.5, 1)$$

$$Simc_{i3}^u = (1, 0.5, 1) \quad Simc_{i4}^u = (0.5, 0.5, 0.5)$$

$$Simc_{i5}^u = (1, 1, 0.5) \quad Simc_{i6}^u = (0.5, 0.5, 1)$$

$$Simc_{i7}^u = (0, 0, 0.5) \quad Simc_{i8}^u = (0.5, 0.5, 0.5)$$

$$Simc_{i9}^u = (0, 0.5, 0.5) \quad Simc_{i10}^u = (0.5, 0.5, 0.5)$$

Paso 2. Construcción de la matriz contextual

$$CONTEXT_u = \begin{pmatrix} 0.5 & 0.5 & 1 \\ 0.5 & 0.5 & 1 \\ 1 & 0.5 & 1 \\ 0.5 & 0.5 & 0.5 \\ 1 & 1 & 0.5 \\ 0.5 & 0.5 & 1 \\ 0 & 0 & 0.5 \\ 0.5 & 0.5 & 0.5 \\ 0 & 0.5 & 0.5 \\ 0.5 & 0.5 & 0.5 \end{pmatrix}$$

Paso 3. Filtrado FBCC.

Los ítems que incumplen la incompatibilidad contextual son los puntos turísticos i_7 e i_9 , ya que

$$\min(simc_{t1}^{7-u}, simc_{t2}^{7-u}, simc_{t3}^{7-u}) = 0 \text{ y } \min(simc_{t1}^{9-u}, simc_{t2}^{9-u}, simc_{t3}^{9-u}) = 0$$

Paso 4. Se calcula el valor C_{avg} para los ocho puntos turísticos restantes:

$$C_{avg} = 7,47$$

Por lo que los ítems i_3 , i_4 e i_6 son también eliminados por no cumplir la condición $C_{i_k} \geq C_{avg}$, seleccionándose¹⁰⁶ las oportunas $V_{i_k}^{select}$ correspondientes a los siguientes turistas, cuyos descriptores turísticos-demográficos suponemos que son:

	d_1	d_2	d_3
i_1	E	E	E
i_2	M	M	Na
i_5	Na	P	Na
i_8	Na	No	M
i_{10}	No	M	M

Tabla 6.10. Descriptores turísticos-demográficos

Paso 5. Se calculan las similitudes Sim_{t-i}^u entre las características del turista que solicita la recomendación y el turista seleccionado para cada ítem que puede recomendarse

$$Sim_{t-i1}^u = (0.5, 0.5, 0.5) \quad Sim_{t-i2}^u = (1, 1, 0)$$

$$Sim_{t-i5}^u = (0, 0, 0) \quad Sim_{t-i8}^u = (0, 0.5, 1)$$

$$Sim_{t-i10}^u = (0.5, 1, 1)$$

donde t_{ik} es el turista que se ha seleccionado por el sistema entre los que han visitado y valorado el punto turístico i_k .

¹⁰⁶ No presentamos la selección de la valoración para cada ítem para simplificar la comprensión del ejemplo, asumiendo que para cada ítem el número de valoraciones tiene que ser alto.

Paso 6. Se calcula el valor de idoneidad de la valoración por parte del turista ti_k del ítem k , por ser un turista más similar al que realiza la petición de recomendación (u), donde q es el cardinal del conjunto de descriptores.

$$idon(ti_k, i_k, u) = \sum_{j=1}^q sim_{d_j}^{i_k}(ti_k - u)$$

$$idon(ti_1, i_1, u) = 1.5$$

$$idon(ti_2, i_2, u) = 2$$

$$idon(ti_5, i_5, u) = 0$$

$$idon(ti_8, i_8, u) = 1$$

$$idon(ti_{10}, i_{10}, u) = 2.5$$

Paso 7. Para realizar el cálculo de predicción y sin pérdida de generalidad, se asume que el valor $V_{i_k}^{select}$ coincide con el valor medio C_{i_k} ¹⁰⁷ que aparece en la tabla 6.8, con lo que los respectivos valores de predicción aplicando la expresión $pred_u^i = idon(ti_i, i, u) \cdot V_{i_k}^{select} + v_i^{conf}$ son:

$$pred_u^{i_1} = 1.5 \cdot 7.25 + 6.4$$

$$pred_u^{i_2} = 2 \cdot 8.48 + 7.6$$

$$pred_u^{i_5} = 0 \cdot 7.8 + 6.1$$

$$pred_u^{i_8} = 1.5 \cdot 8.2 + 8.7$$

$$pred_u^{i_{10}} = 2.5 \cdot 7.9 + 7.7$$

¹⁰⁷ Por definición el valor debe ser superior a C_{i_k} , pero dado que el valor suele ser muy similar, por simplicidad y por no introducir más datos en el ejemplo, se supone que son iguales

	i_1	i_2	i_5	i_8	i_{10}
pred_u^i	17,28	24,56	6,1	21	27,45

Tabla 6.11. Valoración de predicción de interés de los ítems

Por tanto, en este ejemplo el motor de recomendación demográfico recomendaría los puntos turísticos en el siguiente orden:

$$i_{10}, i_2, i_8, i_1, i_5$$

6.4.5. Consideraciones finales

El motor de recomendación presentado en este apartado proporciona novedosas aportaciones respecto a la mayoría de los sistemas de recomendación demográficos existentes (Pazzani, 1999; Vozalis y Margaritis, 2007;; Pérez, 2008):

- a) Permite realizar recomendaciones sin necesidad de tener información de visitas anteriores del turista que solicita la recomendación, lo que soluciona el problema de arranque en frío (cold-start) que tienen muchos sistemas demográficos (Wang y Kao, 2013).
- b) Genera recomendaciones teniendo en cuenta no sólo las valoraciones realizadas por turistas que pertenecen a su mismo grupo demográfico, sino que también tiene influencia el resto de las valoraciones, lo que permite generar en las recomendaciones un efecto sorpresa que normalmente no se produce (Schwab, Kobsa y Koychev, 2001; Fesenmaier et al., 2006).
- c) Utilización de atributos contextuales y de variables lingüísticas.

d) El modelo ha sido diseñado con el objetivo de permitir el efecto sorpresa al no limitar las recomendaciones a los ítems mejor valorados en el grupo demográfico del turista, abriéndolas a ítems valorados positivamente por algún turista con expectativas similares a las suyas. Este aspecto es realmente un elemento innovador que algunos autores consideran oportuno provocar en el usuario con objeto de que las recomendaciones no sean siempre del mismo tipo (Fesenmaier et al., 2006). Al final de esta sección detallamos varias pruebas realizadas de la que se obtienen unas interesantes conclusiones.

Este método podría también tener en cuenta los descriptores que categorizan a los ítems, pero esto conllevaría un exceso de cálculos que se puede evitar. Por otro lado, otra alternativa también correcta sería calcular la valoración media de los ítems considerando únicamente las emitidas en situaciones contextuales similares a las del turista que solicita la recomendación, datos que están disponibles en la estructura del sistema propuesto.

En las pruebas realizadas, se obtienen resultados eficientes en situaciones donde el número de ítems eliminados en la fase de pre-filtrados es grande, siendo pocos los ítems que tendrán que participar en el proceso de recomendación.

A continuación vamos a presentar varias gráficas basadas en unas simulaciones¹⁰⁸ realizadas con diversos números de ítems y diferente

¹⁰⁸ Las diferentes pruebas realizadas las hemos denominado P1, P2, P3, P4 y P5.

número de evaluaciones. Se ha pretendido diferenciar nuestro modelo de un modelo demográfico general en el que el efecto sorpresa no existe. Nuestro modelo lo denominaremos M1 y el modelo demográfico general lo denominaremos M2. El modelo M2 que se ha utilizado se ha implementado teniendo en cuenta únicamente las valoraciones realizadas por los turistas “vecinos”.

En las figuras 6.16, 6.17, 6.18, 6.19 y 6.20 se puede observar que el número de puntos pre-filtrados en la fase cuarta es muy similar en ambos métodos:

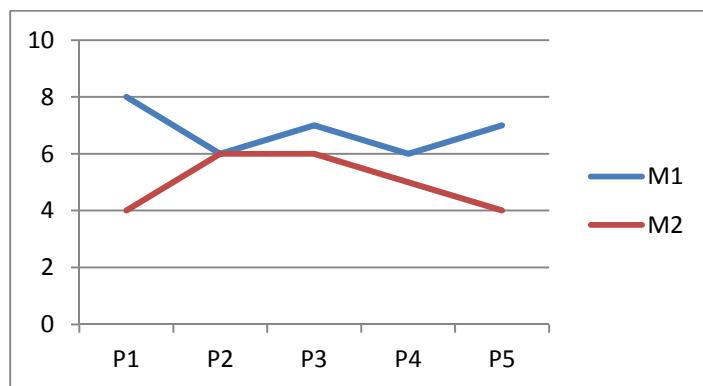


Figura 6.16. Destino con diez ítems y diez calificaciones por ítem

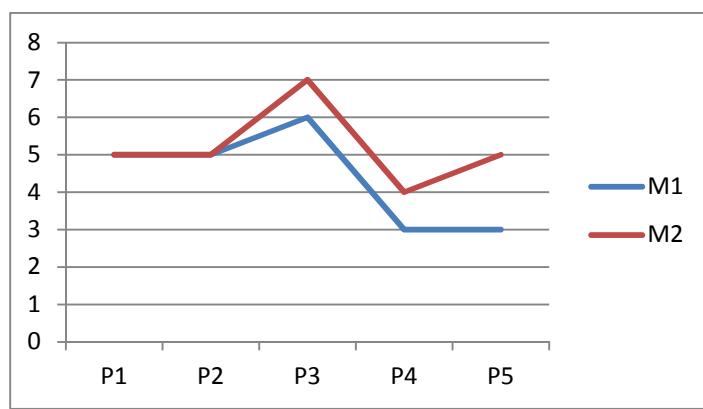


Figura 6.17. Destino con diez ítems y cien calificaciones por ítem

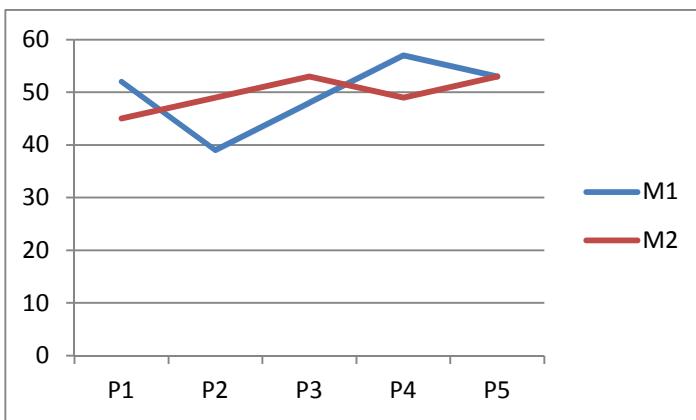


Figura 6.18. Destino con cien ítems y cien calificaciones por ítem

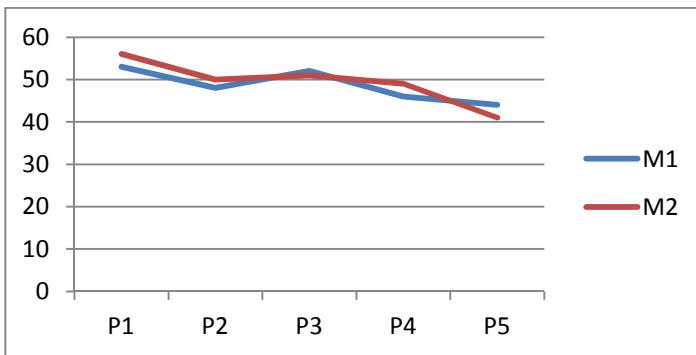


Figura 6.19. Destino con cien ítems y mil calificaciones por ítem

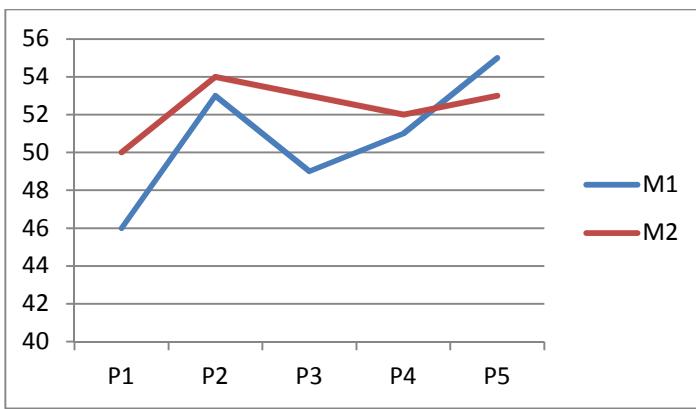


Figura 6.20. Destino con cien ítems y 10000 calificaciones por ítem

Además, para estudiar el grado de sorpresa que puede causar en los turistas las recomendaciones, se han realizado diferentes simulaciones, comprobando las coincidencias entre los cinco primeros ítems que se recomiendan en ambos métodos.

En las figuras 6.21 a 6.26 podemos ver el número de coincidencias entre los primeros cinco ítems en las simulaciones realizadas:

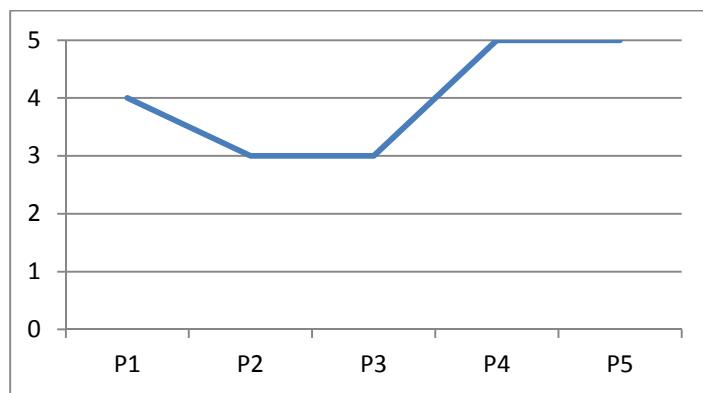


Figura 6.21. Destino con diez ítems y diez calificaciones por ítem

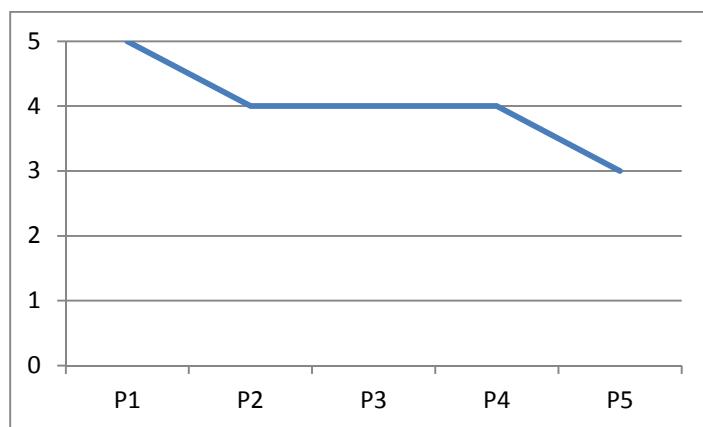


Figura 6.22. Destino con diez ítems y cien calificaciones por ítem

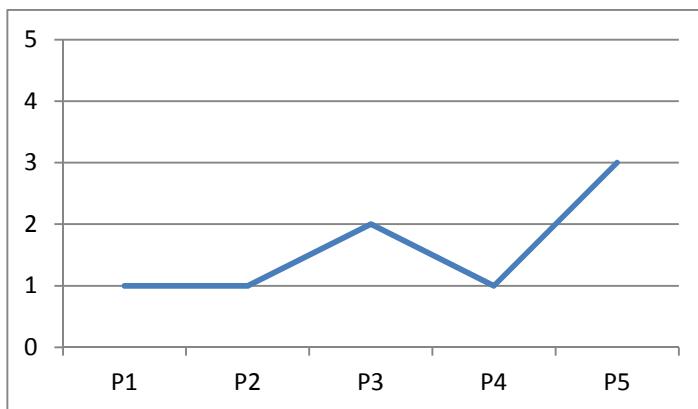


Figura 6.23. Destino con cien ítems y cien calificaciones por ítem

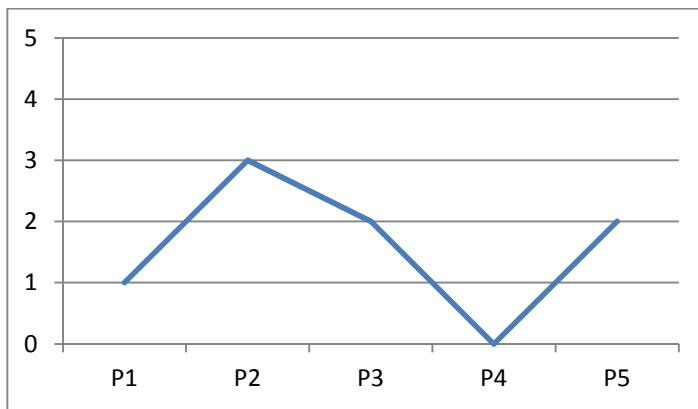


Figura 6.24. Destino con cien ítems y mil calificaciones por ítem

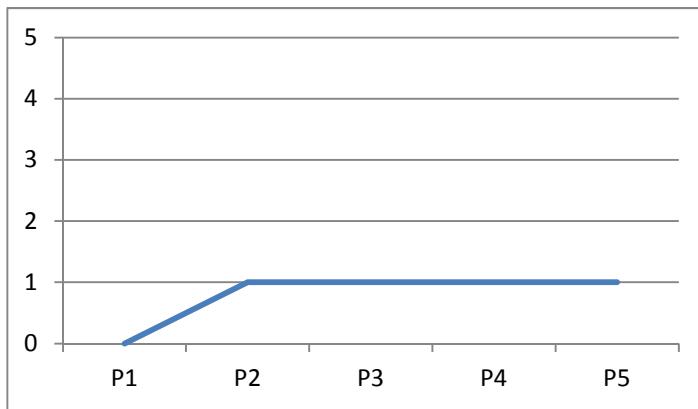


Figura 6.25. Destino con cien ítems y 10000 calificaciones por ítem

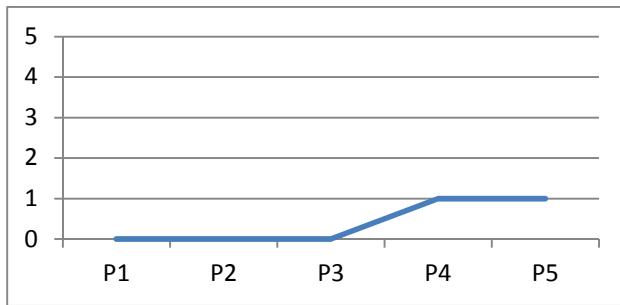


Figura 6.26. Destino con mil ítems y 10000 calificaciones por ítem

Como puede observarse en las figuras anteriores, el factor de sorpresa como era de esperar, aumenta cuando el número de ítems es grande, siendo bastante adecuado en el caso de un número de ítems pequeño, ya que el hecho de que de cinco ítems que visite un turista, uno o dos no coincidan exactamente permitirá que el turista pueda conocer otros ítems bien valorados por otros grupos demográficos.

6.5. Motor de recomendación basado en técnicas colaborativas (SRP^u)

El objetivo de esta sección es describir detalladamente el motor de recomendación colaborativo incluido en nuestro marco de trabajo. Al igual que para los motores de recomendación anteriores, esta sección la dividiremos en diferentes apartados. Uno primero en el que se aporta un resumen de funcionamiento del mismo, para posteriormente realizar la formalización del modelo propuesto y el algoritmo de filtrado y recomendación sobre en el que se basa. Para finalizar se expondrá un ejemplo de su funcionamiento y unas consideraciones finales.

Generalmente este tipo de sistemas de recomendación se basa en la colección completa de ítems valorados por el usuario con anterioridad a la solicitud de recomendación (Eirinaki y Vazirgiannis, 203; Cho y

Kim, 2004; Pérez, 2008). Por tanto, el valor de predicción de interés de un ítem no visitado por un usuario concreto, se calcula como una agregación de las valoraciones de otros sobre dicho ítem (generalmente, se utiliza los k usuarios con valoraciones más parecidos). Por tanto, la descripción de una medida de similaridad es fundamental, para saber si un usuario es similar a un determinado grupo de usuarios.

Al igual que en motores anteriores, se parte de la premisa de que el motor de recomendación debe tener acceso a una base de conocimiento con las valoraciones realizadas por los turistas en sus visitas, así como de las descripciones de las características de los turistas que lo visitan.

6.5.1. Descripción

A continuación se expone el modo de funcionamiento en el que se basa nuestro motor de recomendación basado en técnicas colaborativas. Una de las premisas necesarias para poder ejecutar este motor y que caracteriza a los sistemas de recomendación de este tipo, es que los turistas deben previamente haber visitado diferentes puntos de interés, para que de esta forma se le pueda recomendar aquellas que han sido visitadas por turistas a los que les haya gustado lo mismo que al turista que solicita la recomendación. Cuando éste haya visitado muchas actividades, el cálculo puede ser demasiado complejo, por lo que este método es adecuado utilizarlo cuando el número de puntos visitados en el destino es pequeño.

El funcionamiento del motor, descrito de forma resumida, es el siguiente: partiendo de la información de los ítems ya valorados por el

turista que solicita la recomendación, se seleccionan aquellos otros turistas que han valorado también algunos de esos ítems. A estos turistas se les denominará “amigos”. Posteriormente, se recomendarán otros ítems que hayan sido visitados por los turistas “amigos” y además aquellos ítems que tengan características muy similares a los valorados por estos turistas. Esto último es una aportación interesante del modelo propuesto, ya que los motores de recomendación colaborativos suelen recomendar sólo ítems que ya han sido visitados por turistas que han visitado ítems valorados por el turista que solicita la recomendación.

Previamente se realizará un pre-filtrado contextual siguiendo las pautas ya descritas anteriormente en la sección 6.2. (PFCBI). Una vez aplicado este proceso, se ejecuta un algoritmo basado en técnicas colaborativas que automatiza el proceso de recomendación descrito en el párrafo anterior.

6.5.2. Formalización

Sea I el conjunto de ítems que el sistema detecta a una distancia menor o igual a la indicada por el turista.

$$I = \{i_1, i_2, i_3, \dots, i_n\}$$

Sea U el conjunto de usuarios o turistas que utilizan el motor de recomendación.

$$U = \{u_1, u_2, u_3, \dots, u_m\}$$

Sean i_j e i_k dos ítems o actividades turísticas, se denomina $s(i_j, i_k)$ a la similitud entre los dos ítems. En adelante lo denotaremos como $s_{i_k}^{i_j}$

Sea i_j una actividad o punto turístico y u_k un turista que visita dicho punto o actividad turística. Se denomina $C_{u_k}^{i_j}$ a la valoración o evaluación general de la actividad i_j por parte del turista u_k .

Definición 6.29: Sea i_j un ítem o punto turístico y u_k un turista. Se denomina $P(i_j, u_k)$ al valor de predicción de interés de la actividad i_j para el turista u_k . En adelante lo denotaremos como $P_{u_k}^{i_j}$.

En general calcularemos este valor de predicción para ítems que no han sido visitados por el turista.

Sea I_τ el conjunto índice de características contextuales $\{1, 2, \dots, t\}$. A cada dominio D_i con $i \in I_\tau$ le añadiremos un nuevo elemento denominado ‘indiferente’ y denotado \perp que representará que no se dispone de información acerca de la característica contextual.

Definición 6.30: Se denomina necesidad contextual del turista u , denotado $Nct(u)$, como una tupla formada por la lista de atributos contextuales del turista. Más formalmente:

$$Nct: U \rightarrow \prod_{i \in T} D_i$$

Por tanto

$$Nct(u) = (t_1^u, t_2^u, \dots, t_t^u)$$

Donde t_i^u , es el valor lingüístico del atributo contextual i para el usuario u con $i \in I_\tau$. Esta información es proporcionada por el turista.

Definición 6.31. Se define el concepto de descriptor contextual de la actividad i_k , denotado, $Dsct(i_k)$, como una tupla formada por la lista de valores de atributos contextuales (t). Más formalmente:

$$Dsct: I \rightarrow \prod_{i \in T} D_i$$

El descriptor $Dsct(i_k)$ es por tanto

$$Dsct(i_k) = (t_1^{ik}, t_2^{ik}, \dots, t_t^{ik})$$

Donde t_i^{ik} es el valor lingüístico del atributo contextual i para el ítem i_k , con $i \in I_\tau$. Todos estos valores han de ser proporcionados por los expertos del destino.

Definición 6.32: Se denomina similitud contextual del ítem i y del usuario u en el atributo k (denotada $simc(k,i,u)$), al valor que mide cómo de parecidos son los valores lingüísticos asociados al ítem i y al usuario u . Al valor $simc(k,i,u)$ se le denominará en adelante como $simc_k^{i-u}$. Para calcular este valor se aplicará la medida de posibilidad ya descrita en secciones anteriores, cuya definición es la siguiente:

$$Poss(A, B) = sup_{x \in X} \{min(A(x), B(x))\}$$

Definición 6.33: Se denomina similitud contextual del ítem i y del usuario u , a la lista obtenida a partir de la comparación de los atributos contextuales lingüísticos.

$$Simc(Nct(u), Dsct(i)) = (simc_{t1}^{i-u}, simc_{t2}^{i-u}, \dots, simc_{tr}^{i-u})$$

donde I_τ es el conjunto índice de características contextuales {1, 2, ...r}.

En adelante denominaremos a $Simc(Nct(u), Dsct(i))$ como $Simc_i^u$.

Sea I_c el conjunto índice de características de un ítem $\{1, 2, \dots, c\}$. A cada dominio D_k con $k \in I_c$ le añadiremos un nuevo elemento denominado ‘indiferente’ y denotado \perp que representará que no se dispone de información acerca de la característica.

Definición 6.34. Se define el concepto de descriptor del ítem i_k , denotado, $Ds(i_k)$, como una tupla formada por la lista de valores de características (c). Más formalmente:

$$Ds: I \rightarrow \prod_{k \in C} D_k$$

El descriptor $Ds(i_k)$ es por tanto

$$Ds(i_k) = (c_1^{ik}, c_2^{ik}, \dots, c_c^{ik})$$

Donde c_m^{ik} representa la valoración de la característica m de la actividad i_k con $m \in I_c$. Todos estos valores han de ser proporcionados por los expertos del destino.

Definición 6.35: Se denomina escala experta del atributo j , denotada E^j , al conjunto de posibles valores o términos lingüísticos definidos para calificar el atributo j .

$$E^j = \{e_1^j, e_2^j, \dots, e_z^j\}$$

Por tanto, la valoración de un atributo j por parte del experto del ítem i_k , debe cumplir:

$$c_j^{ik} \in E^j, \quad j \in I_c$$

Para cada atributo o característica se supone que los términos forman una estructura ordenada que cumple:

$$e_i < e_j \Leftrightarrow i < j \text{ y } Neg(e_i) = e_j, \text{ con } j = z - i + 1$$

Para poder realizar operaciones que sirvan para determinar si un ítem es vecino de otro (es decir, que tiene características muy similares) se necesita definir una función de transformación que asigne a cada valor lingüístico un valor numérico igual a la posición que ocupa el término lingüístico. Sea A un conjunto de términos lingüísticos que describe una característica:

$$A = \{e_1, e_2, \dots, e_z\}$$

Se define la función de transformación:

$$F_{TRANS}: A \rightarrow \mathbb{N} / F_{TRANS}(e_i) = i$$

donde i es la posición que ocupa el término lingüístico. Denominaremos rango de una característica j (se denota como $rango(c_j)$) al número de términos lingüísticos que tiene la característica j (cardinal de E^j) .

Definición 6.36: Dos ítems i_1 e i_2 tienen características iguales si todas sus características son idénticas. En este caso diremos que estos ítems son vecinos.

Definición 6.37: Dos ítems i_1 e i_2 con características diferentes son considerados también vecinos, si cumplen la siguiente condición:

$$Ds(i_1) = (c_1^{i_1}, c_2^{i_1}, \dots, c_c^{i_1})$$

$$Ds(i_2) = (c_1^{i_2}, c_2^{i_2}, \dots, c_c^{i_2})$$

$$\sum_{j=1}^c \frac{|F_{TRANS}(c_j^{i1}) - F_{TRANS}(c_j^{i2})|}{rango(c_j)} < umbral$$

El valor del umbral es configurable. En nuestra propuesta se asume, sin pérdida de generalidad, que el umbral es igual a:

$$\sum_{j=1}^c \frac{1}{rango(c_j)}$$

Definición 6.38: Se denomina similitud del ítem i_j y el ítem i_k , al valor obtenido de la siguiente manera:

$$s_{i_k}^{i_j} = \frac{1}{r} \cdot \sum_{m=1}^r sim_{c_m}^{i_j-i_k}$$

donde I_c es el conjunto índice de características de un ítem {1, 2, ..., r}.

Cada $sim_{c_m}^{i_j-i_k}$ se calcula mediante ecuaciones de similaridad, basadas en la medida de posibilidad descrita en secciones anteriores:

$$Poss(A, B) = sup_{x \in X} \{min(A(x), B(x))\}$$

Definición 6.39: Se denomina valor medio de visita del turista u_k , denotado como $\overline{c_{u_k}}$ al valor medio de todas las valoraciones otorgadas por el turista a ítems ya visitados. Así, si el turista ha visitado n ítems en el destino, la expresión para calcularlo sería:

$$\overline{c_{u_k}} = \frac{\sum_{r=1}^n C_{u_k}^{i_r}}{n}$$

Definición 6.40: Se denomina valor medio de visita del ítem i_j , denotado como $\overline{c_{i_j}}$ al valor medio de todas las calificaciones otorgadas por los turistas que han visitado ese ítem. Así, si un ítem ha sido visitado t veces, la expresión sería:

$$\overline{c_{ij}} = \frac{\sum_{r=1}^t C_{ur}^{ij}}{t}$$

Definición 6.41: Se denomina media ajustada de un ítem i_j y un turista u_k que no ha visitado dicho ítem, denotado como $avg_{u_k}^{ij}$, al valor obtenido de la expresión:

$$avg_{u_k}^{ij} = \overline{c_{ij}} + \frac{\sum_{h=1}^n s_{i_h}^{ij} \cdot \max(C_{u_k}^{i_h} - \overline{c_{u_k}}, 0)}{n}$$

Esta expresión pretende aumentar la valoración media a aquellos ítems no visitados por el turista que solicita la recomendación y sean muy similares a ítems visitados por el turista con un valor superior a la media de las calificaciones dadas por este turista. El valor n es el número de visitas que el turista ha realizado y valorado en el destino.

Definición 6.42: Se denomina suma ponderada de un ítem i_j y un turista u_k que no ha visitado dicho ítem, denotado como $pond_{u_k}^{ij}$, al valor obtenido de la expresión:

$$pond_{u_k}^{ij} = \frac{\sum_{h=1}^n s_{i_h}^{ij} \cdot C_{u_k}^{i_h}}{\sum_{h=1}^n C_{u_k}^{i_h}}$$

Este valor calcula la suma de las valoraciones del turista sobre los n ítems visitados, teniendo en cuenta la similaridad respecto al punto sobre el que se desea predecir el interés. El valor de la suma ponderada se emplea en el motor de recomendación para potenciar aquellos ítems que se parezcan mucho a otros que ha sido visitado anteriormente por el turista y que tuvieron una buena valoración.

6.5.3. Proceso de filtrado y recomendación

Para cada ítem detectado por el sistema de realidad aumentada debe realizarse un proceso basado en el siguiente algoritmo:

Paso 0: Filtrado PFCBI (Pre-filtrado contextual basado en implicaciones). Se aplica el proceso de pre-filtrado contextual al conjunto I , obteniéndose un conjunto reducido de ítems (I_R) que será la entrada al siguiente paso del algoritmo de recomendación (véase la sección 6.2).

Paso 1: Para cada ítem i visitado por el turista t se seleccionan todos los turistas que también han valorado ese ítem. A estos turistas se les denomina turistas ‘amigos’.

Paso 2: Se selecciona entre todos aquellos ítems pertenecientes a I_R , aquellos que hayan sido valorado al menos por uno de los turistas ‘amigos’ y que no han sido visitados por el turista que solicita la recomendación. A este conjunto de ítems lo denominaremos $Amig(I)$.

Paso 3: Seleccionar para cada ítem i que pertenezca a $Amig(I)$, aquellos ítems denominados ‘vecinos’ del punto i que también habían sido pre-filtrados en la fase PFCBI. A este conjunto se le denomina $Vec(i)$ y se calcula aplicando las definiciones 6.36 y 6.37

Paso 4: Una vez obtenido el conjunto de ítems $Amig(I) \cup Vec(I)$, se realiza un filtrado contextual. Para todo ítem i que pertenezca al conjunto $Amig(I) \cup Vec(I)$, debe calcularse el valor $Simc_i^u$, siendo u el turista que solicita la recomendación.

Paso 5: Se construye una matriz contextual, formada por las similitudes contextuales de los ítems del conjunto obtenido en el paso anterior (siendo n su cardinal) y las r características de contexto.

$$CONTEXT_u = \begin{pmatrix} Simc_1^u \\ \vdots \\ Simc_n^u \end{pmatrix} = \begin{pmatrix} simc_{t1}^{1-u} & \dots simc_{tj}^{1-u} & \dots & simc_{tr}^{1-u} \\ \vdots & & \vdots & \vdots \\ simc_{t1}^{n-u} & \dots simc_{tj}^{n-u} & \dots & simc_{tr}^{n-u} \end{pmatrix}$$

Paso 6: Filtrado FBCC. Aquellos ítems que incumplan los parámetros contextuales son eliminados de la lista de ítems que pueden ser recomendados. En la siguiente etapa de recomendación sólo se tendrán en cuenta aquellos ítems que verifiquen la siguiente ecuación:

$$\min(simc_{t1}^{i-u}, simc_{t2}^{i-u}, \dots, simc_{tr}^{i-u}) > 0$$

Paso 7: Este paso recoge el proceso de filtrado colaborativo. Se divide en dos etapas que se describen a continuación.

Paso 7.1: Para cada ítem obtenido en la fase anterior se obtiene un valor denominado ‘media ajustada’, calculada como la suma ponderada de las valoraciones realizadas por el turista con un valor superior a la media y su similitud con el punto turístico.

Para ello se utilizan los términos especificados en las definiciones 6.39 y 6.40, donde:

$$\overline{c_{ij}} = \frac{\sum_{r=1}^t C_{ur}^{ij}}{t} \text{ y } \overline{c_{uk}} = \frac{\sum_{r=1}^n C_{ur}^{ij}}{n}$$

Donde n es el número de ítems que ha visitado el turista y t es el número de usuarios¹⁰⁹ que han valorado el ítem del cual se está calculando el valor de media ajustada, definido (definición 6.41) mediante la expresión:

$$\text{avg}_{u_k}^{i_j} = \overline{c_{i_j}} + \frac{\sum_{h=1}^n s_{i_h}^{i_j} \cdot \max(C_{u_k}^{i_h} - \overline{c_{u_k}}, 0)}{n}$$

Paso 7.2: Para cada punto turístico se obtiene un valor denominado ‘suma ponderada’, que resume la evaluación del turista a ítems similares al que se desea predecir (definición 6.42). Este valor se calcula como la suma de las valoraciones del turista sobre los n puntos visitados, teniendo en cuenta la similaridad respecto al ítem sobre el que se desea predecir el interés.

$$\text{pond}_{u_k}^{i_j} = \frac{\sum_{h=1}^n s_{i_h}^{i_j} \cdot C_{u_k}^{i_h}}{\sum_{h=1}^n C_{u_k}^{i_h}}$$

Paso 8: Para cada ítem en el conjunto de salida del paso 6, se calcula un valor de predicción para el usuario que solicita la recomendación de la siguiente forma:

$$P_{u_k}^{i_j} = \text{avg}_{u_k}^{i_j} + \text{pond}_{u_k}^{i_j}$$

Para cada ítem pre-filtrado en el paso 6, deberá realizarse este proceso de cálculo de predicción. Una vez obtenido este valor, se realizará una lista ordenada descendente.

¹⁰⁹ Para los ítems perteneciente a $\text{Amig}(I)$ se utilizan solo las valoraciones de los turistas amigos,

6.5.4. Ejemplo de aplicación del modelo propuesto

En la presente sección se presenta un ejemplo de aplicación del motor de recomendación colaborativo. Supóngase que el sistema detecta que el turista (u_1) que solicita la recomendación ha realizado en el destino previamente visitas a dos ítems i_a e i_b que, a su vez, han sido también valorados por los turistas u_2 y u_3 . Además, se consideran en el ejemplo sólo tres atributos contextuales.

En la siguiente tabla se observa la información sobre las valoraciones¹¹⁰ realizadas por los turistas u_1 , u_2 y u_3 :

Ítem visitado	Valoración turista u_1	Valoración turista u_2	Valoración turista u_3	Valoración media
i_a	10	8	6	8
i_b	8	8	10	8.66

Tabla 6.12. Actividades visitadas por el turista que solicita

recomendación

Se describe a continuación la aplicación de los diferentes pasos del algoritmo propuesto para la recomendación colaborativa.

Paso 0. Al igual que en los motores de recomendación anteriores, se aplica el proceso FPCBI para reducir el número de puntos turísticos a los que aplicar el motor de recomendación.

¹¹⁰ Por simplificación del ejemplo suponemos que los casos donde el turista haya valorado utilizando variables lingüísticas se ha realizado una transformación numérica en una escala de 0 a 10.

Paso 1. Dado que el turista que solicita la recomendación ha visitado dos ítems en el destino, se seleccionan los turistas amigos, que en este caso serían u_2 y u_3 .

Paso 2. A continuación, se deben seleccionar todos aquellos ítems que han pasado el pre-filtrado del paso 0 y que además hayan sido valorado al menos por alguno de los turistas amigos, formando el conjunto de ítems $Amig(I)$. Supongamos que los turistas amigos han visitado las siguientes ítems del destino:

Actividad visitada	Valoración turista u_2	Valoración turista u_3	Valoración turista u_4
i_1	6	10	-
i_2	8	-	-
i_3	8	8	-
i_4	-	6	10
i_5	-	10	10
i_6	8	8	-
i_7	8	-	8

Tabla 6.13. Actividades visitadas por los turistas amigos

En la tabla anterior se ha incluido también la valoración de otro turista denominado u_4 que ha visitado estos puntos. Por simplicidad se han reducido los números de ítems y de turistas.

Paso 3. Se seleccionan para cada punto turístico i perteneciente al conjunto $Amig(I)$, aquellos puntos cercanos a la posición del turista y

filtrado por el sistema PFCBI que sean muy similares, es decir, se obtiene el conjunto $Vec(i)$ de los ítems denominados vecinos del punto i . Para simplificar el problema, supongamos la existencia de tres ítems vecinos que se muestran en la tabla 6.14:

Ítem visitado	Valoración del ítem
i_8	8,3
i_9	7,4
i_{10}	8,9

Tabla 6.14. Actividades vecinas

Paso 4. En el ejemplo propuesto se ha utilizado la escala lingüística que se representa en la figura 6.16, junto a la representación en forma de tupla de las funciones de pertenencia empleadas para representar los atributos contextuales:

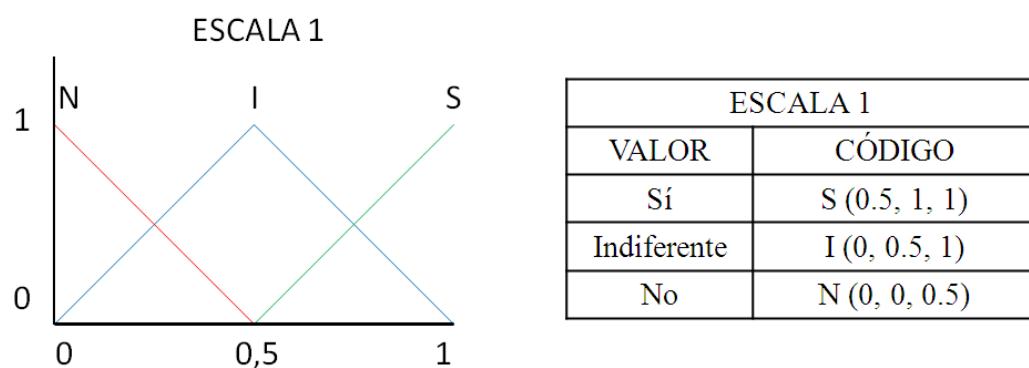


Figura 6.27. Escala lingüística y significado semántico aplicado en el ejemplo

Los atributos contextuales que describen los ítems seleccionados son los siguientes:

	t_1	t_2	t_3
i_1	I	I	I
i_2	I	I	I
i_3	S	I	I
i_4	I	I	S
i_5	S	N	S
i_6	I	I	I
i_7	N	S	N
i_8	I	I	S
i_9	N	I	S
i_{10}	I	I	N

Tabla 6.15. Descripción contextual de los puntos turísticos

Supongamos que el turista que solicita la recomendación está definido por los siguientes parámetros contextuales.

	t_1	t_2	t_3
u_1	S	N	I

Tabla 6.16. Descripción contextual del usuario

Se crea para cada punto turístico la lista:

$$Simc(Nct(u), Dsct(i)) = (simc_{t1}^{i-u}, simc_{t2}^{i-u}, \dots, simc_{tr}^{i-u}) = Simc_i^u$$

$$Simc_{i1}^u = (0.5, 0.5, 1) \quad Simc_{i2}^u = (0.5, 0.5, 1)$$

$$Simc_{i3}^u = (1, 0.5, 1) \quad Simc_{i4}^u = (0.5, 0.5, 0.5)$$

$$Simc_{i5}^u = (1, 1, 0.5) \quad Simc_{i6}^u = (0.5, 0.5, 1)$$

$$Simc_{i7}^u = (0,0,0.5)$$

$$Simc_{i8}^u = (0.5,0.5,0.5)$$

$$Simc_{i9}^u = (0,0.5,0.5)$$

$$Simc_{i10}^u = (0.5,0.5,0.5)$$

Paso 5. Construcción de la matriz contextual

$$CONTEXT_u = \begin{pmatrix} 0.5 & 0.5 & 1 \\ 0.5 & 0.5 & 1 \\ 1 & 0.5 & 1 \\ 0.5 & 0.5 & 0.5 \\ 1 & 1 & 0.5 \\ 0.5 & 0.5 & 1 \\ 0 & 0 & 0.5 \\ 0.5 & 0.5 & 0.5 \\ 0 & 0.5 & 0.5 \\ 0.5 & 0.5 & 0.5 \end{pmatrix}$$

Paso 6. Filtrado FBCC. Los ítems que incumplen la incompatibilidad contextual son i_7 e i_9 , ya que

$$\min(simc_{t1}^{i7-u}, simc_{t2}^{i7-u}, simc_{t3}^{i7-u}) = 0 \text{ y } \min(simc_{t1}^{i9-u}, simc_{t2}^{i9-u}, simc_{t3}^{i9-u}) = 0$$

Paso 7. Comienza el proceso de filtrado colaborativo. Para simplificar, no se presenta el proceso de cálculo de las similitudes (análogo al descrito en el motor de recomendación basado en contenido sin memoria descrito en la sección 6.3.). En la tabla 6.17 se muestran los valores de similaridad de los ítems valorados por turistas “amigos” y otros ítems de características muy similares que hemos denominado “vecinos” respecto a los ítems valorados en el destino por el turista que solicita la petición. La tercera columna de la tabla es la valoración media de cada uno de estos ítems.

	i_a	i_b	\bar{c}_{u_k} ¹¹¹
i_1	0.7	0.9	8
i_2	0.8	1	8
i_3	0.7	0.7	8
i_4	0.5	0.9	6
i_5	1	0.8	10
i_6	0.8	0.9	8
i_8	0.4	0.9	8.3
i_{10}	0.9	0.8	8.9

Tabla 6.17. Tabla de similitudes entre ítems y valor medio de ítem

Como se describió en el apartado anterior, este paso de filtrado colaborativo se divide en dos etapas:

7.1. Para cada ítem que puede ser recomendado se obtiene la ‘media ajustada’, donde:

¹¹¹ Para los ítems desde i_1 hasta i_6 solo se utilizan las valoraciones de los turistas “amigos”

$$\overline{c_{u_1}} = (10+8)/2 = 9$$

Que se obtiene calculando la media de las valoraciones del turista que solicita la recomendación. Posteriormente, se calcula el valor de media ajustada, obteniendo los valores de la tabla 6.18 a partir de la siguiente expresión:

$$avg_{u_k}^{i_j} = \overline{c_{i_j}} + \frac{\sum_{h=1}^n s_{i_h}^{i_j} \cdot \max(C_{u_k}^{i_h} - \overline{c_{u_k}}, 0)}{n}$$

Así, el valor $avg_{u_1}^{i_1}$ se obtendría de la siguiente manera:

$$avg_{u_1}^{i_1} = 8 + \frac{0,7 \cdot 1 + 0,9 \cdot 0}{2} = 8,35$$

$avg_{u_1}^{i_1}$	8,35	$avg_{u_1}^{i_5}$	10,5
$avg_{u_1}^{i_2}$	8,4	$avg_{u_1}^{i_6}$	8,4
$avg_{u_1}^{i_3}$	8,35	$avg_{u_1}^{i_8}$	8,5
$avg_{u_1}^{i_4}$	6,25	$avg_{u_1}^{i_{10}}$	9,35

Tabla 6.18. Tabla de medias ajustadas

Así, el elemento $avg_{u_1}^{i_1}$ se ha obtenido de la siguiente manera:

$$avg_{u_1}^{i_1} = 8 + \frac{0,7 \cdot 1 + 0,9 \cdot 0}{2} = 8,35$$

7.2. Se calcula el valor de suma ponderada para cada ítem (véase la tabla 6.19), aplicando la expresión:

$$\mathbf{pond}_{u_k}^{i_j} = \frac{\sum_{h=1}^n s_{i_h}^{i_j} \cdot C_{u_k}^{i_h}}{\sum_{h=1}^n C_{u_k}^{i_h}}$$

Así, el elemento $\mathbf{pond}_{u_1}^{i_1}$ se obtendría de la siguiente manera:

$$\mathbf{pond}_{u_1}^{i_1} = \frac{0,7 \cdot 10 + 0,9 \cdot 8}{10 + 8} = 0,79$$

$\mathbf{pond}_{u_1}^{i_1}$	0,79	$\mathbf{pond}_{u_1}^{i_5}$	0,91
$\mathbf{pond}_{u_1}^{i_2}$	0,89	$\mathbf{pond}_{u_1}^{i_6}$	0,84
$\mathbf{pond}_{u_1}^{i_3}$	0,7	$\mathbf{pond}_{u_1}^{i_8}$	0,62
$\mathbf{pond}_{u_1}^{i_4}$	0,68	$\mathbf{pond}_{u_1}^{i_{10}}$	0,86

Tabla 6.19. Tabla de sumas ponderadas

Paso 8. Se calcula el valor de predicción de cada ítem para el usuario que solicita la recomendación. Como se describió en el apartado anterior, este valor se calcula la siguiente forma:

$$P_{u_k}^{i_j} = avg_{u_k}^{i_j} + \mathbf{pond}_{u_k}^{i_j}$$

Los valores de predicción se detallan en la siguiente tabla.

$p_{u_1}^{i_1}$	9,14	$p_{u_1}^{i_5}$	11,41
$p_{u_1}^{i_2}$	9,29	$p_{u_1}^{i_6}$	9,24
$p_{u_1}^{i_3}$	9,05	$p_{u_1}^{i_8}$	9,12
$p_{u_1}^{i_4}$	6,93	$p_{u_1}^{i_{10}}$	10,21

Tabla 6.20. Tabla de predicción

Por tanto, el motor de recomendación demográfico que se propone recomendaría los ítems en el siguiente orden (los empates pueden ser resueltos utilizando el parámetro r_i^c):

$$i_5, i_{10}, i_2$$

6.5.5. Consideraciones finales

Tras analizar las características, el comportamiento y la definición de este motor de recomendación, es posible afirmar las siguientes observaciones en cuanto a las situaciones en las que es recomendable su aplicación así como posibles mejoras futuras:

- a) Para utilizar este modelo es necesario que el turista haya visitado anteriormente otros ítems en el destino. Si el turista no ha realizado ninguna visita, la utilización de este motor carece de sentido ya que no se podría obtener una lista de ítems para aplicar el procedimiento propuesto.
- b) Sería una buena opción que el usuario pudiera eliminar de su historial de visitas en el destino aquellas que no tuvieron la

satisfacción esperada para que no participen en el proceso de recomendación, lo que mejoraría la eficiencia de la generación de las recomendaciones al tener que realizar un número de cálculos menores.

- c) Es un motor de recomendación apropiado para usuarios que realizan visitas en grupo, pero que deben expresar las visitas anteriores realizadas por todo el grupo junto, ya que la complejidad de los cálculos que se realizan es menor.
- d) Cuando el número de ítems en un destino es muy grande y también el número de valoraciones, los cálculos requeridos serán muy costosos, lo que quizás pueda provocar tiempos de respuesta excesivos para algunos turistas.

Una posible mejora en el modelo consiste en utilizar las valoraciones contextuales y no generales. En nuestra opinión, este motor es útil sólo cuando es utilizado como complemento a otro motor, pues su principal inconveniente es que tiene un coste computacional elevado para obtener los resultados. El uso previo de otro motor de recomendación como filtro haría que el número de ítems a los que aplicar el modelo sea más pequeño consiguiendo mejorar los tiempos de ejecución.

Finalmente, se resumen las principales ventajas que aporta nuestro modelo respecto a otros motores colaborativos:

- a) Se incorpora información de atributos contextuales.
- b) No se requiere un historial extenso de valoraciones previas del turista que solicita la recomendación, ya que incluso conociendo una sola visita, el motor es capaz de generar recomendaciones.

- c) La mayoría de los motores de recomendación colaborativos usan exclusivamente los ítems visitados por turistas que han valorado los mismos ítems que el que solicita la recomendación, mientras que nuestro motor incluye también otros ítems.

6.6. Motor de recomendación híbrido basado en historial (SHP^u)

La utilización de datos históricos es muy habitual en los sistemas de recomendación colaborativos y basados en contenido (Resnick y Varian, 1997; Maleszka, Mianowska, y Nguyen, 2013; Lee, Kaoli y Huang, 2014). Así, en los motores descritos anteriormente hemos utilizado los datos sobre visitas anteriores del turista que solicita la petición de diferente manera:

- a) Motor basado en contenido con historial: el turista selecciona una visita y el sistema busca ítems que sean similares.
- b) Motor colaborativo: partiendo de ítems valorados por el turista que solicita la recomendación, se identifican turistas que hayan valorado los mismos ítems y se recomiendan otros ítems valorados por estos turistas.

En el caso de los motores de recomendación basados en contenido, el turista tiene que seleccionar o bien las características de los ítems que desea visitar, o bien indicar una visita anterior que le haya agradado para que se busquen ítems similares (Adomavicius y Tuzhilin, 2005; Leiva et al., 2014). Aún, siendo una buena opción, requiere que el turista seleccione el ítem que le interese, además de modificar alguna

característica en caso de que lo considere necesario. Aunque no es un proceso complejo, el modelo que se propone en esta sección, pretende una menor interacción por parte del turista, y que el sistema sea capaz de generar recomendaciones teniendo en cuenta solo las visitas realizadas por el turista pero sin tener que interactuar con el sistema para seleccionar ítems visitados anteriormente o características de ítems que desea visitar.

En esta sección vamos a describir en detalle el citado modelo de recomendación. Como se verá, el modelo utiliza técnicas de sistemas de recomendación basadas en contenido, pero que no tienen en cuenta las características de los ítems que se desea visitar, sino que se basa en las categorías hojas o terminales de la taxonomía (véase la sección 5.6) a la que pertenecen los diferentes ítems que se han visitado.

La técnica que se ha diseñado consiste en generar recomendaciones de ítems que pertenezcan a las categorías que han sido más visitadas por parte del turista y que hayan obtenido mayor valoración. Igual que en los modelos de recomendación presentados en las anteriores secciones, antes de realizar todas estas operaciones debe aplicarse un filtrado previo para eliminar aquellos ítems que no cumplan los atributos contextuales, con objeto de realizar el menor número posible de cálculos.

6.6.1. Descripción

El método de recomendación que se propone en este módulo es similar a otros y consiste en utilizar modelos basados en técnicas bayesianas para la clasificación. Se trata de un método simple pero que da muy buenos resultados (Jensen, 2001; Sarwar et al., 2001; Bueno, 2002;

Blanco, 2007). Como se ha indicado, el algoritmo que se propone se inspira en la teoría probabilística de Bayes, tomando como información de partida las visitas anteriores del turista. El algoritmo genera para cada ítem una estimación de la probabilidad de que el turista desee visitar dicho ítem.

Para poder aplicar el algoritmo es necesario que el turista que realiza la petición de recomendación haya visitado previamente diferentes ítems, y a partir de esta información se le recomendarán aquellos ítems que pertenezcan a categorías que han sido visitadas anteriormente. Así, se generarán mejores resultados cuanto mayor sea el número de ítems valorados por parte del usuario que solicita la recomendación (Pérez, 2008; Guo, Zhang y Thalmann, 2014) De este modo, las actividades que se le recomendarán se parecerán a las que habitualmente está acostumbrado a visitar.

Este modelo de recomendación comienza realizando un estudio sobre las categorías a las que pertenecen los ítems valorados previamente por el turista. El siguiente paso consiste en comprobar cuáles de los ítems cercanos al turista pertenecen a las categorías ya visitadas previamente y cumplen las condiciones contextuales.

6.6.2. Formalización

Sea I el conjunto de ítems situados en el rango de distancia indicado por el turista que solicita la petición de recomendación.

$$I = \{i_1, i_2, i_3, \dots, i_n\}$$

Sea u el usuario que solicita una recomendación. El usuario debe haber valorado previamente diferentes ítems.

Sea I_h el conjunto de ítem visitados por el turista u que solicita la petición de recomendación. Cada punto de interés turístico puede estar asignado en una o varias categorías hojas o terminales C_1, C_2, \dots, C_m (veáse sección 5.6).

Sea i_j un ítem o punto turístico y u_k un turista que no lo ha visitado. Se denomina $P(i_j, u_k)$ al valor de predicción de interés de la actividad i_j por parte del turista u_k . En adelante lo denotaremos como $P_{u_k}^{i_j}$.

Definición 6.43: Se define la probabilidad histórica (denotado $P(C_j^u)$) como la probabilidad de que el turista u visite un ítem de categoría C_j . Esta probabilidad es el cociente entre el número de visitas realizadas por el turista u que pertenezcan a la categoría C_j y el número de visitas totales realizadas por el mismo.

Definición 6.44: Se define la probabilidad histórica normalizada de un turista u y una categoría C_i (denotada por $P_{histn}(C_i^u)$), al valor calculado de la siguiente forma:

$$P_{histn}(C_i^u) = P(C_i^u) / \sum_{k=1}^n P(C_k^u)$$

Siendo n el número de categorías visitadas por el turista. El objetivo de este cálculo es conseguir que la suma de todas las probabilidades históricas normalizadas sumen el valor unidad.

Definición 6.45: Se define el peso normalizado de la categoría c en un ítem i (denotado $w(c,i)$), al valor calculado como sigue:

$$w(c,i) = \frac{k_c^i}{\sum_{j=1}^z k_j^i}$$

donde k_c^i indica el grado de pertenencia del ítem i a esa categoría c (definición 5.6) y z es el número de categorías a las que está asociada ese ítem i (véase la sección 5.6). En adelante lo denotaremos como w_c^i

Definición 6.46: Se define la probabilidad normalizada adaptada al peso de una categoría i para un turista u (denotada $P_w(C_i^u)$) a:

$$P_w(C_i^u) = \frac{\sum_{k=1}^n w(C_i, i_k)}{n}$$

donde n , es el número de visitas realizadas por el turista u . Este valor es mucho más adecuado a la realidad de un turista, ya que se da mayor importancia a las visitas de ítems donde una determinada categoría es predominante o tiene más peso.

Definición 6.47: Se define la probabilidad de predicción de interés de un ítem i_k para el turista u de la siguiente forma:

$$P_u^{i_k} = P(i_k, u) = \sum_{j=1}^n k_{C_j}^{i_k} \cdot P_w(C_j^u) \cdot P_{histn}(C_j^u)$$

Sea I_τ el conjunto índice de características contextuales $\{1, 2, \dots, t\}$. A cada dominio D_i con $i \in I_\tau$ se le añade un nuevo elemento denominado ‘indiferente’ y denotado \perp que representará que no se dispone de información acerca de la característica contextual.

Definición 6.48: Se define la necesidad contextual del turista u , denotada $Nct(u)$, como una tupla formada por la lista de valores de atributos contextuales (t) del turista. Más formalmente:

$$Nct: U \rightarrow \prod_{i \in T} D_i$$

La necesidad contextual $Nct(u)$ es por tanto

$$Nct(u) = (t_1^u, t_2^u, \dots, t_t^u)$$

Donde t_i^u , es el valor lingüístico del atributo contextual i para el usuario u con $i \in I_{\mathcal{T}}$. Esta información es proporcionada por el turista.

Definición 6.49. Se define el concepto de descriptor contextual del ítem i_k , denotado, $Dsct(i_k)$, como una tupla formada por la lista de valores de atributos contextuales (t). Más formalmente:

$$Dsct: I \rightarrow \prod_{i \in T} D_i$$

El descriptor $Dsct(i_k)$ es por tanto

$$Dsct(i_k) = (t_1^{ik}, t_2^{ik}, \dots, t_t^{ik})$$

Donde t_i^{ik} es el valor lingüístico del atributo contextual i para la actividad i_k , con $i \in I_{\mathcal{T}}$. Todos estos valores han de ser proporcionados por los expertos del destino.

Definición 6.50: Se define la similitud contextual del atributo k del ítem i y del usuario u ($simc(k,i,u)$), al valor que mide cómo de parecidos son los valores lingüísticos asociados al ítem i y al usuario u . Al valor

$\text{simc}(k, i, u)$, se le denotará en adelante como simc_k^{i-u} . Para calcular este valor se aplicará la medida de posibilidad ya descrita en secciones anteriores, y que, dada dos funciones de pertenencia A y B se define:

$$\text{Poss}(A, B) = \sup_{x \in X} \{\min(A(x), B(x))\}$$

Definición 6.51: Se denomina similitud contextual del ítem i y del usuario u , a la lista obtenida a partir de la comparación de los atributos contextuales lingüísticos.

$$\text{Simc}(Nct(u), Dsct(i)) = (\text{simc}_{t1}^{i-u}, \text{simc}_{t2}^{i-u}, \dots, \text{simc}_{tr}^{i-u})$$

donde I_r es el conjunto índice de características contextuales {1, 2, ...r}.

En adelante denotaremos a $\text{Simc}(Nct(u), Dsct(i))$ como Simc_i^u .

6.6.3. Proceso de filtrado y recomendación

En esta sección se describe detalladamente el proceso de recomendación que genera las recomendaciones basadas en el historial del turista. El algoritmo de recomendación consiste en los siguientes pasos:

Paso 0: Filtrado PFCBI (Pre-filtrado contextual basado en implicaciones). Se aplica a los ítems cercanos al turista una fase de pre-filtrado contextual (véase la sección 6.2), obteniéndose un conjunto reducido de ítems (I_R) que será la entrada al siguiente paso del algoritmo de recomendación.

Paso 1: Se realiza un nuevo filtrado contextual de los ítems obtenidos en el paso anterior. Para todo ítem $i \in I_R$ que el sistema puede recomendar debe calcularse el valor Simc_i^u

Paso 2: Se construye una matriz contextual, formada por los valores de similitud contextual de los ítems del conjunto generado en el paso anterior (cardinalidad n) y las r características de contexto.

$$CONTEXT_u = \begin{pmatrix} Simc_1^u \\ \vdots \\ Simc_n^u \end{pmatrix} = \begin{pmatrix} simc_{t1}^{1-u} & \dots & simc_{tj}^{1-u} & \dots & simc_{tr}^{1-u} \\ \vdots & & \vdots & & \vdots \\ simc_{t1}^{n-u} & \dots & simc_{tj}^{n-u} & \dots & simc_{tr}^{n-u} \end{pmatrix}$$

Paso 3: Filtrado FBCC. Aquellos ítems que incumplan los parámetros contextuales son eliminados de la lista de ítems que pueden ser recomendados. En otras palabras, un ítem $i \in I_R$ sólo pasará a la siguiente etapa de si verifica que:

$$\min(simc_{t1}^{i-u}, simc_{t2}^{i-u}, \dots, simc_{tr}^{i-u}) > 0$$

A este nuevo subconjunto de ítems filtrados le denominaremos I_F , y se cumple que $I_F \subseteq I_R \subseteq I$.

Paso 4: Tomando en consideración los ítems pertenecientes a I_h , se realiza el cálculo de la probabilidad histórica $P(C_j^u)$ de que el turista u que solicita la recomendación visite un ítem de una determinada categoría C_i . Este valor se calcula para cada categoría hoja de los ítems que haya visitado el turista.

Paso 5: Los valores obtenidos en el paso anterior deben normalizarse, obteniendo para cada categoría la probabilidad histórica normalizada $P_{histn}(C_i^u)$.

Paso 6: Se calcula el peso normalizado de categoría (véase la definición 6.45) para cada ítem que haya sido visitado por el turista.

Paso 7: Se calcula para cada categoría la probabilidad normalizada adaptada al peso (véase definición 6.46).

Paso 8: El sistema debe calcular para todos los ítems que pueden recomendarse el valor de predicción de interés de dicho ítem para el turista. Para ello se tendrá en cuenta el valor k_c^i de los ítems que pueden recomendarse, obteniendo el valor de predicción utilizando la expresión descrita en la definición 6.47:

$$P_u^{i_k} = P(i_k, u) = \sum_{j=1}^n k_{C_j}^{i_k} \cdot P_w(C_j^u) \cdot P_{histn}(C_j^u)$$

La recomendación consistirá en la lista de ítems ordenada de forma decreciente según su valor de predicción.

6.6.4. Ejemplo de aplicación del modelo propuesto

Pasos 0 a 3. Se realiza el pre-filtrado PFCBI y posteriormente un filtrado contextual de forma idéntica a la presentada en los motores anteriores. Para simplificar el ejemplo se supone que solamente pasan el proceso de pre-filtrado y filtrado contextual tres ítems (i_1 , i_2 e i_3).

Paso 4. Cálculo de probabilidad histórica $P(C_i^u)$ del turista que solicita la recomendación con respecto a las diferentes categorías de ítems. En el ejemplo se supone que el turista u ha visitado anteriormente cinco ítems pertenecientes a cuatro categorías hojas:

Turista	Categoría 1	Categoría 2	Categoría 3	Categoría 4
Ítem 1	X	X		
Ítem 2		X		X
Ítem 3			X	
Ítem 4	X	X	X	
Ítem 5			X	X

Tabla 6.21. Visitas realizadas por turista que solicita recomendación

El valor ‘X’ en la tabla indica que ese ítem ha sido visitado, valorado y está relacionado con esa categoría. De esta manera obtendríamos siguiendo la definición 6.43¹¹²:

$$P(C_1^u) = 0,4$$

$$P(C_2^u) = 0,6$$

$$P(C_3^u) = 0,6$$

$$P(C_4^u) = 0,4$$

Paso 5. Cálculo de probabilidad histórica normalizada. Los valores obtenidos en el apartado anterior son normalizados para que la suma de todas las probabilidades sea 1, obteniendo el valor descrito en la definición 6.44, que se calcula de la siguiente manera:

$$P_{histn}(C_i^u) = P(C_i^u) / \sum_{k=1}^n P(C_k^u)$$

En el ejemplo se obtienen los siguientes valores:

$$P_{histn}(C_1^u) \frac{0,4}{0,4 + 0,6 + 0,6 + 0,4} = 0,2$$

$$P_{histn}(C_2^u) \frac{0,6}{0,4 + 0,6 + 0,6 + 0,4} = 0,3$$

$$P_{histn}(C_3^u) \frac{0,6}{0,4 + 0,6 + 0,6 + 0,4} = 0,3$$

$$P_{histn}(C_4^u) \frac{0,4}{0,4 + 0,6 + 0,6 + 0,4} = 0,2$$

Paso 6. Cálculo de pesos normalizados de la categoría para cada ítem visitado. Como se indicó anteriormente, estos pesos son calculados aplicando la fórmula:

¹¹² Denominaremos C_1 a la categoría 1, C_2 a la categoría 2', etc.

$$w(c, i) = \frac{k_c^i}{\sum_{j=1}^z k_j^i}$$

Los valores obtenidos en el paso anterior ya permitirían realizar un estudio probabilístico adecuado, pero es posible obtener resultados más fiables, dado que el sistema almacena para cada ítem un grado de pertenencia o interés del punto respecto a cada categoría (definición 5.6).

Supóngase un ítem i que pertenece a dos categorías hojas (C_1 y C_2), y que en la primera tiene un grado de pertenencia máxima (valor unidad) y en la otra categoría tiene un valor 0,5. Los valores de peso normalizado para ese ítem en cada una de esas categorías se obtendrían de la siguiente manera:

$$w(C_1, i) = \frac{1}{1 + 0,5} = 0,66$$

$$w(C_2, i) = \frac{0,5}{1 + 0,5} = 0,33$$

Supóngase que se obtienen los valores de peso normalizado en las diferentes categorías de las distintas visitas realizadas por el turista que solicita la recomendación. Para simplificar el ejemplo, no se presenta la forma de calcularlos, indicando en la siguiente tabla los valores resultantes:

Turista	Categoría 1	Categoría 2	Categoría 3	Categoría 4
Ítem 1	0.6	0.4	0	0
Ítem 2		0.9		0.1
Ítem 3			1	
Ítem 4	0,1	0,1	0,8	
Ítem 5			0,8	0,2

Tabla 6.22. Tabla de pesos normalizados para cada visita

Como puede observarse, el ítem 1 está relacionado con la categoría 1 y categoría 2, pero la pertenencia de este ítem a la categoría 1 es más relevante que a la categoría 2.

Paso 7. Cálculo de la probabilidad normalizada adaptada. Utilizando la expresión:

$$P_w(C_i^u) = \frac{\sum_{k=1}^n w(C_i, i_k)}{n}$$

se obtendría en el ejemplo que:

$$P_w(C_1^u) = \frac{0,6 + 0,1}{5} = \frac{0,7}{5} = 0,14$$

$$P_w(C_2^u) = \frac{0,4 + 0,9 + 0,1}{5} = \frac{1,4}{5} = 0,28$$

$$P_w(C_3^u) = \frac{1 + 0,8 + 0,8}{5} = \frac{2,6}{5} = 0,52$$

$$P_w(C_4^u) = \frac{0,1 + 0,2}{5} = \frac{0,3}{5} = 0,06$$

Paso 8. Cálculo de predicción. El sistema debe calcular para todos los ítems que pueden recomendarse la probabilidad que dicho punto sea de

interés al turista. Supongamos que los tres ítems que se han obtenido tras el paso 3 (i_1 , i_2 e i_3) están catalogados dentro de cinco categorías hoja. Partamos de los siguientes valores:

Puntos de interés.	Categoría 1	Categoría 2	Categoría 3	Categoría 4	Categoría 5
i_1	0,3	0,1	0	0,7	0
i_2	0	0	0,8	0,2	0
i_3	0,4	0,3	0	0	0,3

Tabla 6.23. Tabla de valores k_c^i

Con estos datos se pueden obtener las probabilidades de interés de cada uno de los ítems para el turista mediante la expresión:

$$P_u^{i_k} = P(i_k, u) = \sum_{j=1}^n k_{C_j}^{i_k} \cdot P_w(C_j^u) \cdot P_{histn}(C_j^u)$$

Resultando por tanto:

$$\begin{aligned} P_u^{i_1} &= 0,3 \cdot 0,14 \cdot 0,2 + 0,1 \cdot 0,28 \cdot 0,3 + 0 \cdot 0,52 \cdot 0,3 + 0,7 \cdot 0,06 \cdot 0,2 + 0 \\ &= 0,0252 \end{aligned}$$

$$P_u^{i_2} = 0 \cdot 0,14 \cdot 0,2 + 0 \cdot 0,28 \cdot 0,3 + 0,8 \cdot 0,52 \cdot 0,3 + 0,2 \cdot 0,06 \cdot 0,2 = 0,1272$$

$$P_u^{i_3} = 0,4 \cdot 0,14 \cdot 0,2 + 0,3 \cdot 0,28 \cdot 0,3 + 0 \cdot 0,52 \cdot 0,3 + 0 \cdot 0,06 \cdot 0,2 + 0 = 0,0364$$

Como ninguno de los ítems visitados por el turista pertenecía a la categoría 5, la probabilidad que se le asigna por defecto a esa categoría es cero.

Según los valores obtenidos para el turista en cuestión, se le debe recomendar el ítem i_2 como elemento turístico que será más probable que cumpla sus preferencias.

6.6.5. Consideraciones finales

El motor de recomendación presentado en esta sección tiene como principal ventaja su baja complejidad algorítmica. Una posible mejora al modelo propuesto sería considerar las valoraciones que tienen los ítems que pueden recomendarse por parte de los turistas que lo han visitado (definición 5.9), pudiendo obtener el valor de probabilidad de interés de cada uno de los ítems de la siguiente manera:

$$P_u^{i_k} = P(i_k, u) = \sum_{j=1}^n k_{C_j}^{i_k} \cdot P_w(C_j^u) \cdot P_{histn}(C_j^u) \cdot r_{i_k}^{C_j}$$

Otra posible modificación al motor propuesto sería permitir que el turista que solicita la recomendación pueda si lo desea eliminar algunos ítems visitados por el propio turista para que no participe en la generación de las recomendaciones. Una ventaja que aportaría este cambio es que la recomendación se asemejará más al interés del turista, pero, por el contrario tiene el inconveniente de que el turista debe interaccionar con el sistema para indicarle esta información. Esto sería contrario al objetivo que se proponía en este modelo, que era conseguir que las recomendaciones se generen con la menor interacción posible.

El motor de recomendación presentado en este apartado es aplicable principalmente en las siguientes situaciones:

- a) El usuario que realiza peticiones de recomendación debe obligatoriamente haber realizado visitas anteriormente.

- b) Usuarios que realizan visitas en grupo, deben expresar las visitas que hayan sido realizadas por el grupo.

6.7. Motor de recomendación grupal

En esta sección se va a describir el motor de recomendación grupal incorporado al marco de trabajo propuesto. Este tipo de motores es especialmente importante en el ámbito turístico, ya que la mayoría de las visitas que se realizan son en compañía. Por este motivo, es muy interesante que un destino turístico disponga de recursos para proporcionar recomendaciones a sus visitantes, pero además, resulta indispensable la posibilidad de generar resultados teniendo en cuenta las preferencias de los turistas que forman el grupo.

Se trata de una línea de investigación en auge que en estos últimos años (Jameson y Smyth, 2007; Cabrerizo, 2008; Chen, Cheng y Chuang, 2008; Piliponyte, 2012). Uno de los principales problemas que se debe resolver es la gran cantidad de información que debe gestionarse (McCarthy, 2002), así como el tratamiento de aspectos relacionados con la personalidad de los individuos que forman el grupo (Leiva et al., 2014).

Cuando se van a realizar recomendaciones grupales en un entorno orientado al turismo pueden seguirse dos estrategias (Leiva et al., 2014):

- a) Estudiar las preferencias individuales de todos los miembros del grupo, generando recomendaciones para cada uno de los miembros del grupo, para posteriormente combinar los

resultados, eligiendo aquellos ítems comunes a todas las recomendaciones ordenados de mayor a menor valor de predicción.

- b) Crear un nuevo turista virtual, que estará descrito mediante el valor medio de las diferentes características de todos los miembros del grupo. Las recomendaciones para el grupo serán las que se generen para ese turista virtual.

Nuestra propuesta combina las dos aproximaciones anteriores, creando un turista virtual para los atributos contextuales y así construir un primer filtro sobre el conjunto de ítems, y generando a partir de los ítems resultantes las recomendaciones individuales. Posteriormente, mediante funciones de agregación en las que pueden tenerse en cuenta parámetros relacionados con la personalidad de los miembros del grupo, se obtendría la lista ordenada de ítems a recomendar.

Como puede deducirse, para generar la recomendación grupal, se necesita utilizar un motor de recomendación individual. En la presente sección se ha optado (sin perder la generalidad) por emplear el motor de recomendación basado en contenido sin memoria. Dado que este motor de recomendación se ha descrito en la sección 6.3, se utilizarán varios de los conceptos allí presentados sin repetir aquí gran parte de las definiciones.

6.7.1. Descripción

Dado que vamos a utilizar un motor de recomendación basado en contenido sin memoria, se necesita considerar datos de diferentes fuentes de información y de diverso tipo. Las fuentes serán los turistas que expresan sus preferencias y los puntos de interés que son descritos por

expertos del destino. Para que el modelo tenga utilidad real para un destino turístico, las descripciones de los diferentes ítems tienen que ser realizadas por personal cualificado.

Los turistas deberán expresar utilizando valoraciones lingüísticas, tanto las características de lo que le gustaría visitar, así como los descriptores turísticos-demográficos que indican el tipo de turista que es, así como sus atributos contextuales que en este caso son comunes al resto de los miembros que componen el grupo.

Como se describió en la sección 6.3, los motores de recomendación basados en contenido se encargan de buscar ítems que tengan descripciones similares a las necesidades solicitadas. En muchos casos las propiedades o atributos que describen un ítem pueden estar descritos por un valor numérico, pero en nuestro caso se ha optado por facilitar a los turistas descripciones más flexibles, utilizando las escalas lingüísticas adecuadas (véase sección 6.3.3). Para ello, los distintos puntos de interés que pueden recomendarse deben estar descritos por medio de atributos relacionados con aspectos cualitativos, cuyo dominio estará descrito en un dominio lingüístico. Dado que las diferentes características pueden ser muy diversas, es adecuado el uso de diferentes escalas.

El punto inicial del proceso de recomendación grupal que se propone es la recopilación de información del perfil de todos los turistas que forman el grupo. En concreto, y tal como se detalló anteriormente, debe recopilarse las características más importantes de los ítems o puntos de interés turísticos que les gustaría visitar, información sobre el tipo de

turista de cada uno de los miembros del grupo, así como información contextual relevante.

A continuación, el sistema identifica una serie de ítems cercanos a la localización de los miembros del grupo. Después se ejecuta una primera fase de pre-filtrado contextual (PFCBI), descrita en la sección 6.2 y aplicada en los motores de recomendación anteriores. Esta fase tiene como objeto reducir el número de ítems con los que se trabaja en la etapa posterior, descartando aquellos ítems que son contextualmente incompatibles con el grupo.

Una vez generado el conjunto de ítems resultante de la fase de filtrado, el motor de recomendación grupal utilizará la información recopilada acerca de las necesidades o preferencias que tiene cada turista del grupo. Debe destacarse que en este motor de recomendación se propone el uso de información de las necesidades actuales del turista. De esta forma, éste podrá expresar necesidades o preferencias que pueden no ser las habituales o incluso podrá indicar el tipo de ítem que desea visitar, como puede ser sólo restaurantes o museos por ejemplo.

Como ya se destacó en la sección 6.3, una diferencia destacable entre nuestra propuesta y otros motores de recomendación basados en contenido reside en el tipo de información sobre el perfil del turista, ya que la mayoría de ellos se utiliza exclusivamente información histórica (de visitas anteriores), mientras que en el motor que estamos describiendo esa fuente de información puede ser sustituida por una necesidad especificada explícitamente por el turista antes de la solicitud de la

recomendación, complementada por información contextual y sobre el tipo de turista que es.

A partir de los perfiles de los turistas, de las descripciones de los ítems a recomendar y qué tipo de turistas suelen visitarlos se realiza un proceso de filtrado. En este proceso se calcula la similitud entre el perfil del turista individual y los diferentes ítems a recomendar. En este punto es importante recordar la posibilidad de que las descripciones puedan estar en diferentes escalas, lo que hará necesario definir operaciones para poder comparar las valoraciones de las características de los puntos y del perfil del turista. Una vez obtenida esta información, el sistema genera una lista ordenada, teniendo en cuenta el grado de similitud de los ítems con las necesidades especificadas por el turista y sus propias características. Cuando se tiene toda esta información de recomendación para cada uno de los miembros que forman el grupo, se generará una única lista para el grupo, que estará formada por la unión de los ítems o puntos de interés incluidos en las recomendaciones individuales, y ordenada de acuerdo a la posición obtenida por los ítems en cada una de éstas. La ordenación podrá modularse teniendo en cuenta tanto valoraciones realizadas por otros turistas que han visitado previamente el ítem, como el grado de flexibilidad (Leiva et al., 2014) de los diferentes miembros del grupo. El concepto de flexibilidad está relacionado con el grado de comprensión que puede tener un turista al visitar algo que no le gusta pero que si es del agrado de sus acompañantes.

Para simplificar el ejemplo que propondremos en secciones posteriores, se asume que tanto el experto como el turista utilizan la misma escala (cualitativa) para valorar las características. Por otro lado, se considera que las valoraciones de los ítems son numéricas. En la figura 6.28 se ilustra el funcionamiento del sistema cuando un grupo formado por tres turistas solicitan una recomendación.

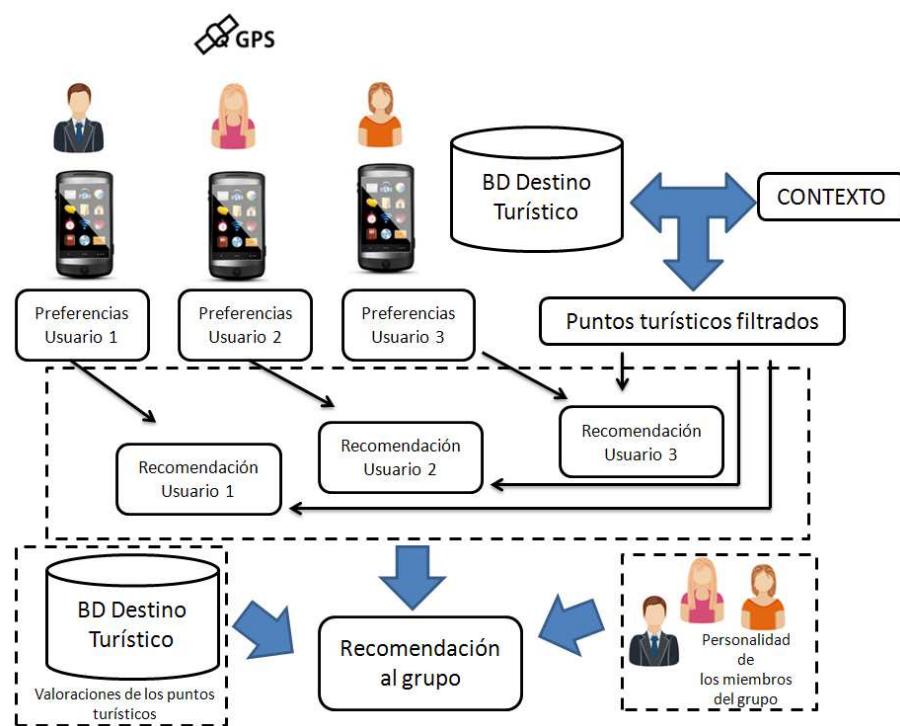


Figura 6.28. Esquema recomendación grupal

6.7.2. Formalización de la representación de la información y medidas de comparación

En este apartado se detallan las definiciones de los conceptos sobre ítems y preferencias de turistas, así como de diferentes operaciones basadas en enfoques lingüísticos difusos necesarias en este motor de recomendación. Algunas son prácticamente idénticas a las descritas en

la sección 6.3.4 y 6.3.5, pero consideramos adecuado repetirlas para recordarlas.

Sea I el conjunto de ítems localizados en el rango de distancia indicado por los turistas.

$$I = \{i_1, i_2, i_3, \dots, i_n\}$$

Sea I_T el conjunto índice de características contextuales $\{1, 2, \dots, t\}$, I_D el conjunto índice de descriptores turísticos y demográficos $\{1, 2, \dots, d\}$, e I_C el conjunto índice de características de un ítem $\{1, 2, \dots, c\}$.

A cada dominio D_i con $i \in I_T$ se le añade un nuevo elemento denominado ‘indiferente’ y denotado \perp que representará la ausencia de información acerca de la característica c_i . Análogamente, se añade el elemento \perp a cada dominio D_j con $j \in I_D$ y para cada D_k con $k \in I_C$.

Definición 6.52. Se define el concepto de descriptor de la actividad i_k , denotado, $Ds(i_k)$, como una tupla formada por la concatenación de las listas de valores de atributos contextuales (t), valores de descriptores turísticos y demográficos (d) y de valores de características (c). Más formalmente:

$$Ds: I \rightarrow \prod_{i \in T} D_i \times \prod_{j \in D} D_j \times \prod_{k \in C} D_k$$

El descriptor $Ds(i_k)$ es por tanto

$$Ds(i_k) = (t_1^{ik}, t_2^{ik}, \dots, t_t^{ik}, d_1^{ik}, d_2^{ik}, \dots, d_d^{ik}, c_1^{ik}, c_2^{ik}, \dots, c_c^{ik})$$

Donde t_i^{ik} es el valor del atributo contextual i para la actividad i_k , con $i \in I_{\mathcal{T}}$, d_j^{ik} es el valor del descriptor j para la actividad i_k con $j \in I_{\mathcal{D}}$ y c_m^{ik} representa la valoración de la característica m de la actividad i_k con $m \in I_{\mathcal{C}}$. Todos estos valores han de ser proporcionados por los expertos del destino.

Definición 6.53: Se define necesidad del turista u , denotada $N(u)$, como una tupla formada por la concatenación de las lista de valores de atributos contextuales (t) del grupo (g) del que forma parte el usuario u , de valores de descriptores turísticos-demográficos (d) que lo definen y de valores de las características de los ítems que describen las preferencias “actuales” (p) del turista, Más formalmente:

$$N: U \rightarrow \prod_{i \in T} D_i \times \prod_{j \in D} D_j \times \prod_{k \in C} D_k$$

Las necesidad $N(u)$ es por tanto

$$N(u) = (t_1^g, t_2^g, \dots, t_t^g, d_1^u, d_2^u, \dots, d_d^u, p_1^u, p_2^u, \dots, p_c^u)$$

Donde t_i^g con $i \in I_{\mathcal{T}}$ es el valor del atributo contextual i para el grupo g al que pertenece el usuario u, con $i \in I_{\mathcal{T}}$, d_j^u con $j \in I_{\mathcal{D}}$ es el valor del descriptor j para el turista u y p_m^u con $m \in I_{\mathcal{C}}$ representa la valoración de la característica m por parte del turista u. Esta información es proporcionada por el turista.

Definición 6.54: Dado $j \in I_{\mathcal{C}}$, se denomina escala experta del atributo j, denotada E^j , como el conjunto de posibles valores o términos lingüísticos definidos para calificar el atributo j.

$$E^j = \{e_1^j, e_2^j, \dots, e_z^j\}$$

Por tanto, la valoración de un atributo j por parte del experto del punto o actividad i_k , debe cumplir:

$$c_j^{ik} \in E^j, \quad j \in I_c$$

De forma análoga se valorarán los atributos contextuales y los descriptores turísticos-demográficos.

Definición 6.55: Dado $j \in I_c$, se define la escala de usuario del atributo j , denotada EU^j , al conjunto de posibles valores o términos lingüísticos definidos para calificar el atributo j .

$$EU^j = \{eu_1^j, eu_2^j, \dots, eu_z^j\}$$

Por tanto, la valoración del atributo j por parte del usuario u , debe cumplir:

$$p_j^u \in EU^j, \quad j \in I_c$$

En adelante se considera que $E^j = EU^j$, para de esta forma simplificar el modelo propuesto. De forma análoga se definiría para los atributos contextuales y los descriptores turísticos-demográficos

Se denomina filtrado basado en contenido contextual grupal FBCCG, al proceso mediante el cual se seleccionan del conjunto de ítems de I que han sido pre-filtrados mediante el motor PFCBI (véase sección 6.2), aquellos que cumplen los atributos contextuales de todos los miembros del grupo.

Se denomina filtrado basado en contenido sin memoria grupal FBSMG, al proceso mediante el cual se selecciona entre cuáles de los ítems que cumplen los atributos contextuales del grupo, aquellos que deben ser recomendados teniendo en cuenta el perfil de todos los turistas que forman el grupo.

Para cada ítem debe realizarse un proceso de medida de similitud que compare los descriptores y características de los puntos con las preferencias del turista.

Definición 6.56: Se define el concepto de similitud del atributo¹¹³ k del ítem i y de la preferencia del usuario u en ese atributo k (denotada $\text{sim}(k,i,u)$), al valor que mide cómo de parecidos son los valores lingüísticos asociados al ítem i y al usuario u . Al valor $\text{sim}(k,i,u)$, se le denota alternativamente como sim_k^{i-u} . El valor de similitud se calcula aplicando la medida de posibilidad descrita en la sección 6.3.3.

Definición 6.57: Se define el concepto de similitud contextual del atributo k del ítem o actividad turística i y del contexto actual del grupo g en ese atributo k (denotada $\text{simc}(k,i,g)$), al valor que mide cómo de parecidos son los valores lingüísticos asociados al ítem i y al grupo g . Al valor $\text{simc}(k,i,g)$, se le denota alternativamente como simc_k^{i-g} . El valor de similitud se calcula aplicando la medida de posibilidad de funciones de pertenencia descrita en la sección 6.3.3.

Definición 6.58: Se define la similitud contextual del ítem o actividad turística i y del grupo g al que pertenece el usuario u (denotada como

¹¹³ Descriptor turístico-demográfico o característica

$Simc(N(g), Ds(i))$, como la lista obtenida a partir de la comparación de los atributos contextuales lingüísticos.

$$Simc(N(g), Ds(i)) = (simc_{t1}^{i-g}, simc_{t2}^{i-g}, \dots, simc_{tr}^{i-g})$$

donde I_{τ} es el conjunto índice de características contextuales {1, 2, ...r}.

En adelante denotaremos a $Simc(N(g), Ds(i))$ como $Simc_i^g$.

Definición 6.59: Se define la similitud del ítem i y del usuario u (denotada como $Sim(N(u), Ds(i))$), como la lista obtenida a partir de la comparación de los respectivos descriptores y características no contextuales.

$$Sim(N(u), Ds(i)) = (sim_{d1}^{i-u}, sim_{d2}^{i-u}, \dots, sim_{dq}^{i-u}, sim_{c1}^{i-u}, sim_{c2}^{i-u}, \dots, sim_{cm}^{i-u})$$

donde $I_{\mathcal{D}}$ es el conjunto índice de descriptores turísticos y demográficos {1, 2,q}, e I_c el conjunto índice de características de un ítem {1, 2,m}. Como ya se ha defnid cada sim_k^{i-u} se calcula mediante ecuaciones de similaridad, basadas en la medida de posibilidad. En adelante se denotará a $Sim(N(u), Ds(i))$ como Sim_i^u .

Definición 6.60: Se define el valor de predicción de interés para el usuario u del ítem i y se denota como $pred(i, u)$ al valor obtenido mediante la siguiente expresión:

$$pred(i, u) = \sum_{j=1}^q sim_{d_j}^{i-u} + \sum_{m=1}^c sim_{c_m}^{i-u}$$

con $j \in I_{\mathcal{D}}$ y $m \in I_c$. En adelante se denotará el valor de predicción de usuario como $pred_u^i$.

Definición 6.61: Se define el valor de predicción de interés para el grupo g del ítem i y se denota como $\text{pred}(i, g)$ al valor obtenido mediante la siguiente expresión:

$$\text{pred}(i, g) = \sum_{k=1}^n \text{pred}(i, u_k)$$

siendo n el número de turistas que forman el grupo g . En adelante se denotará el valor de predicción de grupo como pred_g^i .

6.7.3. Proceso de recomendación

En este apartado se describe de forma detallada el proceso que aplica este modelo de recomendación para obtener una lista de ítems ordenadas según las preferencias del grupo que solicita la recomendación. Los pasos a seguir son los siguientes¹¹⁴:

Paso 0: Filtrado PFCBI (Pre-filtrado contextual basado en implicaciones). A los ítems cercanos (I) se les aplica una fase de pre-filtrado contextual (véase la sección 6.2), obteniéndose un conjunto reducido de ítems (I_R) que serán la entrada al siguiente paso del algoritmo de recomendación.

Paso 1: Para todo ítem o punto de interés perteneciente a I_R debe calcularse el valor Simc_i^g

Paso 2: Se construye una matriz contextual, formada por los valores de similitud de los ítems pertenecientes a I_R y las r características de contexto. Se supone que el cardinal de es $I_R n$.

¹¹⁴ Algunos de los pasos coinciden con el resto de motores, pero con objeto de facilitar la lectura y comprensión, se ha optado por describirlos nuevamente. De la misma manera también se han incluido en los pasos del algoritmo algunos conceptos adicionales que para una mejor comprensión era adecuada contextualizarla en la propia descripción del paso.

$$CONTEXT_g = \begin{pmatrix} Simc_1^g \\ \vdots \\ Simc_n^g \end{pmatrix} = \begin{pmatrix} simc_{t1}^{1-g} & \dots & simc_{tj}^{1-g} & \dots & simc_{tr}^{1-g} \\ \vdots & & \vdots & & \vdots \\ simc_{t1}^{n-g} & \dots & simc_{tj}^{n-g} & \dots & simc_{tr}^{n-g} \end{pmatrix}$$

Paso 3: Filtrado FBCCG (Filtrado basado en características contextuales grupales). Aquellos ítems o puntos de interés que incumplan los parámetros de incompatibilidad contextual son eliminados de la lista de ítems que pueden ser recomendados. En otras palabras, un ítem $i \in I_R$ pasará a la siguiente etapa de recomendación sólo si satisface la siguiente ecuación:

$$\min(simc_{t1}^{i-g}, simc_{t2}^{i-g}, \dots, simc_{tr}^{i-g}) > 0$$

A este nuevo subconjunto de ítems filtrados le denominaremos I_F , y cumple que $I_F \subseteq I_R \subseteq I$.

Paso 4: Para todo ítem $i \in I_F$ y para todo usuario u del grupo g debe calcularse el valor Sim_i^u . Es importante recordar que en este valor no participan los valores contextuales, solo los descriptores turísticos-demográficos y características no contextuales.

Paso 5: Para cada ítem $i \in I_F$ se calcula su valor de predicción de interés para el grupo (denotado como pred_g^i), basándonos en el valor de predicción de interés para cada usuario del grupo. La forma de calcular este valor ha sido descrita en la definición 6.61.

Paso 6: Ordenar la lista de ítems pertenecientes a I_F teniendo en cuenta el valor pred_g^i . En el caso de obtener para cada ítem valores diferentes entre sí, ya habría finalizado el proceso de recomendación. Si existen

ítems con valores idénticos, se debe realizar un proceso de desempate que se describe en los dos siguientes pasos.

Paso 7: Construir una matriz de interés grupal para aquellos ítems que tienen un mismo valor pred_g^i . Para ello se calcula el valor $\text{pref}(i_a, i_b, g)$ que mide el número de características en las que los diferentes usuarios del grupo prefieren el ítem i_a al ítem i_b . En adelante a este valor lo denotaremos como $\text{pref}_{i_a i_b}$, y la forma de calcularlos es:

$$\text{pref}(i_a, i_b, g) = \text{pref}_{i_a i_b} = \sum_{k=1}^n [\sum_{j=1}^q \max(\text{sim}_{d_j}^{i_a-u} - \text{sim}_{d_j}^{i_b-u}, 0) + \sum_{m=1}^c \max(\text{sim}_{c_m}^{i_a-u} - \text{sim}_{c_m}^{i_b-u}, 0)]$$

Donde n es el número de turistas que forman el grupo, con $j \in I_\mathcal{D}$ y $m \in I_\mathcal{C}$. Se calculan estos valores para los z puntos empatados con el mismo valor de predicción que pueden recomendarse, y se construye una matriz de interés que nos mide las preferencias entre las diferentes ítems por parte del grupo g .

$$\text{PREF}_g = \begin{pmatrix} \text{pref}_{11} & \dots & \text{pref}_{1j} & \dots & \text{pref}_{1z} \\ \vdots & & \vdots & & \vdots \\ \text{pref}_{z1} & \dots & \text{pref}_{zj} & \dots & \text{pref}_{zz} \end{pmatrix}$$

Una vez calculada la matriz, se deben ordenar los diferentes puntos según el grado de satisfacción de los turistas del grupo. Para ello se calcula el nivel de preferencia estricto para cada punto o actividad turística. La matriz resultante se denomina $\text{PREF}_g^{\text{estricta}}$, donde cada elemento está calculado de la siguiente manera (Wang, 2005; Martínez et al., 2008):

$$\text{pref}_{ab}^{\text{estricta}} = \max(\text{pref}_{ab} - \text{pref}_{ba}, 0)$$

De este modo se obtiene para cada ítem un valor de grado de desempate (GD), calculado de la siguiente forma:

$$GD_i = \sum_{j=1}^m pref_{ij} + pref_{ij}^{estricta}$$

Paso 8: Reordenar la lista de ítems. En caso de que aún existan ítems con el mismo valor de preferencia, se multiplican sus grados de desempate GD_i respectivos por el valor r_i^c (véase la definición 5.9) que tenga un grado de pertenencia k (véase la definición 5.6) mayor de entre las diferentes categorías en la que está catalogado el ítem i . En caso de que varios ítems tengan el mismo grado de pertenencia k , se seleccionará el que tenga un valor r_i^c mayor. Si aún así, existieran elementos con idéntico valor se informaría que tienen el mismo grado de interés. El objetivo de este criterio de ordenación es dar mayor valor a aquellos ítems que tengan una mayor valoración por parte de los turistas que lo han visitado.

Como se describió en la sección 6.3 una alternativa a tener en cuenta en caso de tener varios ítems con el mismo valor de predicción es la aplicación posterior (técnica de hibridación en cascada) de un segundo motor de recomendación, para ordenar de esta forma los ítems con igual valor de preferencia.

6.7.4. Proceso de mejoras en la recomendación para grupos.

Una posible mejora podría obtenerse comparando los valores $pred(i, g)$ con las valoraciones realizadas por otros turistas que han visitado ese ítem, así como por un factor de dominancia en la

personalidad de los miembros del grupo (Leiva et al, 2012). En concreto, los pasos a aplicar serían los siguientes:

- Maximizar la satisfacción media. Para cada ítem se calcula el valor medio $\text{pred}(i, g)$ de la siguiente forma (donde n es el número de turistas):

$$\text{pred}(i, g) = \frac{1}{n} \cdot \sum_{k=1}^n \text{pred}(i, u_k)$$

En esta mejora se eliminarían del conjunto de ítems a recomendar aquellos que verifiquen que más de la mitad de los miembros del grupo tienen un valor de predicción inferior al valor medio $\text{pred}(i, g)$.

- Minimizar la carencia. En esta aproximación, se penaliza a aquellos ítems i que tiene un valor $\text{pred}(i, u)$ muy bajo para un usuario u del grupo g . Más concretamente, a todo ítem se le aplica el siguiente ajuste en su valor de predicción:

$$\text{Si } \exists u \in g \text{ / } \text{pred}(i, u) = 0, \text{ entonces } \text{pred}(i, g) = \text{pred}(i, g) - \mu$$

donde μ es el valor de penalización. Este grado de penalización debe minimizarse con el objetivo de asegurar un grado de justicia. El motivo es que se considera más adecuado recomendar ítems que satisfagan a todos los miembros del grupo de una forma similar antes que aquellos que satisfagan más a unos pocos.

Otra opción para refinar el modelo de recomendación grupal consiste en considerar los grados de personalidad de cada uno de los turistas que

forman el grupo. Esta aproximación puede formalizarse de la siguiente manera:

Sea w_i el grado de dominancia del turista i , que indica si el turista tiene una personalidad dominante o no. Así, a una persona dominante, le asignaremos el valor 1, mientras que una persona nada dominante le asignaremos el valor 0,5.

Definición 6.62: Se denomina valor de grado de dominancia normalizado del turista i y se denota como W_i al valor obtenido utilizando la siguiente expresión:

$$W_i = \frac{w_i}{\sum_{j=1}^n w_j}$$

Por tanto se puede calcular el valor de predicción para el grupo con la siguiente expresión:

$$pred(i, g) = \sum_{k=1}^n pred(i, u_k) \cdot W_i$$

donde n es el número de usuarios que forman el grupo.

6.7.5. Ejemplo de aplicación del modelo propuesto

En esta sección se presenta un ejemplo de aplicación del motor de recomendación grupal.

En el ejemplo se considera que el grupo está formado por tres turistas y que el sistema de PFCBI (Paso 0) ha pre-filtrado un total de diez ítems. Al igual que en los motores de recomendación anteriores, por sencillez se ha reducido el número de puntos turísticos, así como el número de características, descriptores y atributos contextuales. Así,

cada ítem vendrá descrito por tres atributos contextuales, tres descriptores turísticos y cinco características.

Sean (t_1, t_2, t_3) los atributos contextuales, (d_1, d_2, d_3) los descriptores turísticos y $(c_1, c_2, c_3, c_4, c_5)$ las características de los ítems. Cada una de estas características están valoradas por los expertos mediante varias escalas lingüísticas.

En el ejemplo propuesto se han utilizado cuatro escalas triangulares, que se representan en la figura 6.29 junto a las tablas con los números difusos (representación en forma de tuplas de las funciones de pertenencia) de dichas escalas (Figura 6.30):

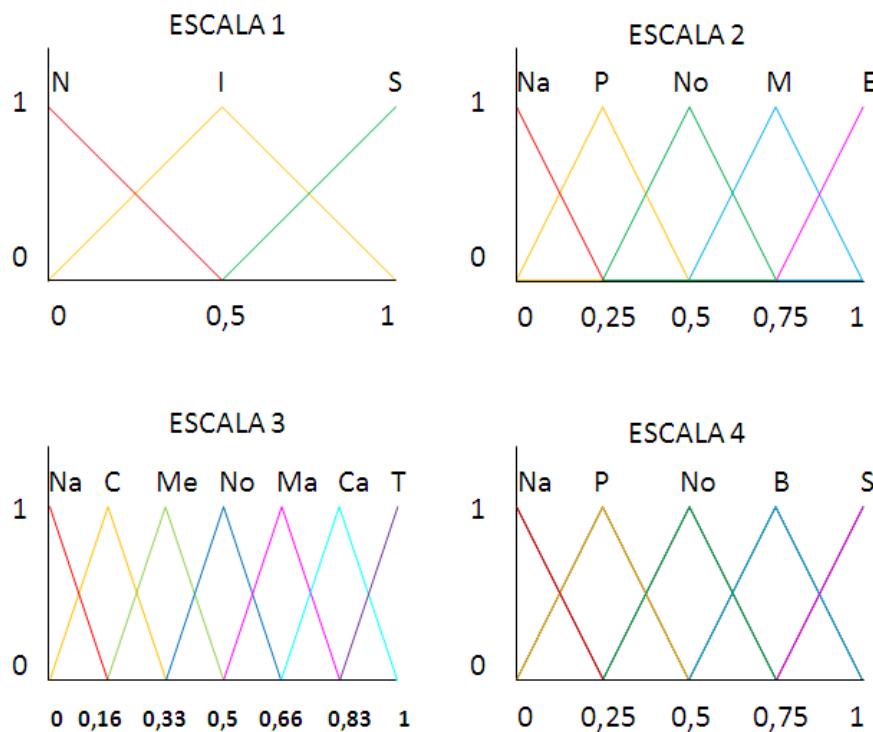


Figura 6.29. Escalas lingüísticas utilizadas en el ejemplo

ESCALA 1		ESCALA 2	
VALOR	CÓDIGO	VALOR	CÓDIGO
Sí	S (0.5, 1, 1)	Excelente	E (0.75, 1, 1)
Indiferente	I (0, 0.5, 1)	Mucho	M (0.5, 0.75, 1)
No	N (0, 0, 0.5)	Normal	No(0.25,0.5,0.75)

ESCALA 3		ESCALA 4	
VALOR	CÓDIGO	VALOR	CÓDIGO
Totalmente	T (0.83, 1, 1)	Sí	S (0.75, 1, 1)
Casi todo	Ca (0.66, 0.83, 1)	Bastante	B (0.5, 0.75, 1)
Más de lo normal	Ma (0.5, 0.66, 0.83)	Normal	No(0.25,0.5,0.75)
Normal	No(0.33,0.5,0.66)	Poco	P (0, 0.25, 0.5)
Menos de lo normal	Me(0.16,0.33,0.5)	Nada	Na (0, 0, 0.25)
Casi nada	C(0, 0.16, 0.33)		
Nada	Na (0, 0, 0.16)		

Figura 6.30. Significado semántico de las escalas utilizadas en el ejemplo

Las características que describen los puntos turísticos son los siguientes:

	t_1	t_2	t_3	d_1	d_2	d_3	c_1	c_2	c_3	c_4	c_5
i_1	I	I	I	S	S	S	Me	S	T	No	No
i_2	I	I	I	B	B	Na	C	S	Ca	No	E
i_3	S	I	I	No	P	No	C	N	Ca	Me	No
i_4	I	I	S	P	No	No	No	I	Ma	Na	No
i_5	S	N	S	Na	P	Na	Ma	S	No	C	P
i_6	I	I	I	No	S	No	T	I	No	Me	No
i_7	N	S	N	B	S	No	Me	N	Ca	No	P
i_8	I	I	S	Na	No	B	T	N	T	T	Na
i_9	N	I	S	P	No	B	Me	I	Ca	No	E
i_{10}	I	I	N	No	B	B	Na	I	No	Ma	P

Tabla 6.24. Lista de puntos turísticos para recomendar.

Concretamente, para los atributos contextuales se ha utilizado la escala 1 y para los descriptores se ha utilizado la escala 4. Además, para los atributos de características de los diferentes puntos turísticos se ha empleado la escala 3 para las características primera, tercera y cuarta, la escala 1 para la segunda característica y la escala 2 para la quinta.

El sistema requiere a cada usuario del grupo expresar sus preferencias y necesidades respecto a los puntos turísticos que desea le sean recomendados. Los atributos contextuales de cada turista como puede observarse en el ejemplo deben ser comunes, pues estos atributos corresponden al grupo.

	t_1	t_2	t_3	d_1	d_2	d_3	c_1	c_2	c_3	c_4	c_5
u_1	S	N	I	No	B	B	Na	I	No	No	Na
u_2	S	N	I	B	B	B	C	I	Me	Ma	P
u_3	S	N	I	No	S	S	Na	I	No	Ma	P

Tabla 6.25. Lista de usuarios que forman un grupo

Se detalla a continuación la aplicación del resto de los pasos del algoritmo.

Paso 1. Para cada ítem se calcula $Simc_i^g$

$$Simc_1^g = (0.5, 0.5, 1)$$

$$Simc_2^g = (0.5, 0.5, 1)$$

$$Simc_3^g = (1, 0.5, 1)$$

$$Simc_4^g = (0.5, 0.5, 0.5)$$

$$Simc_5^g = (1, 1, 0.5)$$

$$Simc_6^g = (0.5, 0.5, 1)$$

$$Simc_7^g = (0, 0, 0.5)$$

$$Simc_8^g = (0.5, 0.5, 0.5)$$

$$Simc_9^g = (0, 0.5, 0.5)$$

$$Simc_{10}^g = (0.5, 0.5, 0.5)$$

Paso 2. Construcción de la matriz contextual

$$CONTEXT_g = \begin{pmatrix} 0.5 & 0.5 & 1 \\ 0.5 & 0.5 & 1 \\ 1 & 0.5 & 1 \\ 0.5 & 0.5 & 0.5 \\ 1 & 1 & 0.5 \\ 0.5 & 0.5 & 1 \\ 0 & 0 & 0.5 \\ 0.5 & 0.5 & 0.5 \\ 0 & 0.5 & 0.5 \\ 0.5 & 0.5 & 0.5 \end{pmatrix}$$

Paso 3. Filtrado FBCCG.

Los puntos que incumplen la incompatibilidad contextual son los puntos turísticos i_7 e i_9 , ya que

$$\min(simc_{t1}^{7-g}, simc_{t2}^{7-g}, simc_{t3}^{7-g}) = 0$$

$$\min(simc_{t1}^{9-g}, simc_{t2}^{9-g}, simc_{t3}^{9-g}) = 0$$

Paso 4. Para los ocho puntos ítems filtrados en el paso anterior se calcula el valor Sim_i^u para todos los usuarios del grupo.

Así, para el usuario 1 los valores obtenidos serían¹¹⁵:

$$Sim_{i1}^{u1} = (0, 0.5, 0.5 | 0, 0.5, 0, 1, 0)$$

$$Sim_{i2}^{u1} = (0.5, 1, 0 | 0.5, 0.5, 0, 1, 0)$$

$$Sim_{i3}^{u1} = (1, 0, 0.5 | 0.5, 0.5, 0, 0.5, 0)$$

$$Sim_{i4}^{u1} = (0.5, 0.5, 0.5 | 0.1, 0.5, 0, 0)$$

$$Sim_{i5}^{u1} = (0, 0, 0 | 0, 0.5, 1, 0, 0.5)$$

$$Sim_{i6}^{u1} = (1, 0.5, 0.5 | 0.1, 1, 0.5, 0)$$

$$Sim_{i8}^{u1} = (0, 0.5, 1 | 0, 0.5, 0, 0, 1)$$

$$Sim_{i10}^{u1} = (1, 1, 1 | 1, 1, 1, 0.5, 0.5)$$

Para el usuario 2:

¹¹⁵ Se ha utilizado una barra vertical para separar los descriptores de los atributos no contextuales para facilitar la comprensión del ejemplo.

$$\begin{array}{ll}
 Sim_{i1}^{u2} = (0.5, 0.5, 0.5|0.5, 0.5, 0, 0.5, 0.5) & Sim_{i2}^{u2} = (1, 1, 0|1, 0.5, 0, 0.5, 0) \\
 Sim_{i3}^{u2} = (0.5, 0, 0.5|1, 0.5, 0, 0, 0.5) & Sim_{i4}^{u2} = (0, 0.5, 0.5|0.1, 0, 0, 0.5) \\
 Sim_{i5}^{u2} = (0, 0, 0|0, 0.5, 0.5, 0, 1) & Sim_{i6}^{u2} = (0.5, 0.5, 0.5|0.1, 0.5, 0, 0.5) \\
 Sim_{i8}^{u2} = (0, 0.5, 1|0, 0.5, 0, 0, 0.5) & Sim_{i10}^{u2} = (0.5, 1, 1|0.5, 1, 0.5, 1, 1)
 \end{array}$$

Y para el usuario 3:

$$\begin{array}{ll}
 Sim_{i1}^{u3} = (0, 1, 1|0, 0.5, 0, 0.5, 0.5) & Sim_{i2}^{u3} = (0.5, 0.5, 0|0.5, 0.5, 0, 0.5, 0) \\
 Sim_{i3}^{u3} = (1, 0, 0|0.5, 0.5, 0, 0, 0.5) & Sim_{i4}^{u3} = (0.5, 0, 0|0.1, 0.5, 0, 0.5) \\
 Sim_{i5}^{u3} = (0, 0, 0|0, 0.5, 1, 0, 1) & Sim_{i6}^{u3} = (1, 1, 0|0.1, 1, 0, 0.5) \\
 Sim_{i8}^{u3} = (0, 0, 0.5|0, 0.5, 0, 0, 0.5) & Sim_{i10}^{u3} = (1, 0.5, 0.5|1, 1, 1, 1, 1)
 \end{array}$$

Paso 5: Para cada ítem se calcula el valor de predicción de interés para el grupo:

$$\begin{aligned}
 pred(i_1, g) &= \sum_{k=1}^3 pred(i_1, u_k) = pred(i_1, u_1) + pred(i_1, u_2) + pred(i_1, u_3) \\
 &= 2.5 + 3 + 3.5 = 9.0
 \end{aligned}$$

$$\begin{aligned}
 pred(i_2, g) &= \sum_{k=1}^3 pred(i_2, u_k) = pred(i_2, u_1) + pred(i_2, u_2) + pred(i_2, u_3) \\
 &= 3 + 4 + 2.5 = 9.5
 \end{aligned}$$

$$\begin{aligned}
 pred(i_3, g) &= \sum_{k=1}^3 pred(i_3, u_k) = pred(i_3, u_1) + pred(i_3, u_2) + pred(i_3, u_3) \\
 &= 3 + 3 + 2.5 = 8.5
 \end{aligned}$$

$$\begin{aligned}
 pred(i_4, g) &= \sum_{k=1}^3 pred(i_4, u_k) = pred(i_4, u_1) + pred(i_4, u_2) + pred(i_4, u_3) \\
 &= 3 + 2.5 + 2.5 = 8
 \end{aligned}$$

$$\begin{aligned} \text{pred}(i_5, g) &= \sum_{k=1}^3 \text{pred}(i_5, u_k) = \text{pred}(i_5, u_1) + \text{pred}(i_5, u_2) + \text{pred}(i_5, u_3) \\ &= 2 + 2 + 2.5 = 6.5 \end{aligned}$$

$$\begin{aligned} \text{pred}(i_6, g) &= \sum_{k=1}^3 \text{pred}(i_6, u_k) = \text{pred}(i_6, u_1) + \text{pred}(i_6, u_2) + \text{pred}(i_6, u_3) \\ &= 4.5 + 3.5 + 4.5 = 12.5 \end{aligned}$$

$$\begin{aligned} \text{pred}(i_8, g) &= \sum_{k=1}^3 \text{pred}(i_8, u_k) = \text{pred}(i_8, u_1) + \text{pred}(i_8, u_2) + \text{pred}(i_8, u_3) \\ &= 3 + 2.5 + 1.5 = 7 \end{aligned}$$

$$\begin{aligned} \text{pred}(i_{10}, g) &= \sum_{k=1}^3 \text{pred}(i_{10}, u_k) = \text{pred}(i_{10}, u_1) + \text{pred}(i_{10}, u_2) \\ &\quad + \text{pred}(i_{10}, u_3) = 7 + 6.5 + 7 = 20.5 \end{aligned}$$

Paso 6: Ordenar la lista de ítems teniendo en cuenta el valor pred_g^i . En este caso como los valores de predicción de cada ítem son diferentes, el algoritmo finalizaría recomendando al grupo la siguiente lista ordenada de ítems:

$$i_{10}, i_6, i_2, i_1, i_3, i_4, i_8, i_5$$

En el caso de que hubiese ítems con el mismo valor de predicción, se aplicarían los dos últimos pasos del algoritmo, en un proceso similar al que se mostró en el ejemplo de la sección 6.3.6.

6.7.6. Consideraciones finales

En esta sección se ha presentado un modelo de recomendación grupal basado en contenido sin memoria, que utiliza parámetros contextuales. Esto constituye una importante aportación respecto a otros modelos existentes (McCarthy, 2002, Jameson y Smyth, 2007:

Cabrero, 2008; Chen, Cheng y Chuang, 2008; Piliponyte, 2012). Otra aportación importante de nuestra propuesta es la inclusión de parámetros relacionados con la personalidad de los miembros del grupo (Cabrero, 2008; Leiva et al., 2014).

En esta sección se ha mostrado como construir un motor de recomendación grupal como una modificación del motor de recomendación individual basado en contenido sin memoria. Aplicando el mismo esquema, se podría diseñar un motor de recomendación grupal partiendo de cualquiera de los restantes motores de recomendación individuales presentados.

6.8. Creación de rutas. Trazabilidad turística.

En esta sección se presenta una propuesta para aplicar las cadenas de Markov a un proceso de generación de rutas, que partiendo de los ítems que hayan sido recomendados a un turista por un motor de recomendación. El modelo propuesto está fuertemente relacionado con el concepto de trazabilidad turística del destino. El concepto genérico de trazabilidad (Walderhaug et al., 2009) se relaciona con el estudio de los procesos que permiten conocer la evolución de un determinado producto. En el ámbito turístico la trazabilidad permite conocer la evolución de un ítem a partir de las visitas que realizan los diferentes turistas. Esta información puede resultarnos de mucha utilidad para la generación de rutas personalizadas a partir del análisis de las secuencias de visitas más frecuentes (Leiva, Guevara y Rossi, 2012).

Esta sección se ha organizado en diferentes apartados que describen las ideas básicas que se necesitan para comprender las cadenas de Markov, su aplicación en un destino turístico para la generación de rutas y el estudio de la trazabilidad turística mediante diferentes parámetros. Se terminará con un ejemplo en el que se podrá comprobar su utilidad y funcionamiento.

6.8.1. Introducción

Cuando un turista desea visitar varios ítems o puntos de interés, no siempre lo hace siguiendo estrictamente el orden de preferencia o al orden sugerido por un motor de recomendación. Normalmente el turista visita en primer lugar aquel punto turístico que cree que le va a gustar más y después va visitando otros ítems, teniendo en cuenta diferentes factores que van cambiando en cada momento, como la distancia, tiempo, cansancio, etc.

Actualmente, se observa un cambio en el comportamiento de los turistas, que cada vez tienden a sustituir paquetes estándar a otras opciones más personalizadas (Rodríguez, Molina y Caballero, 2010). Por tanto, los destinos deben adaptarse a dicho cambio ofreciendo herramientas que consideren las preferencias de los turistas (Valls, 2003).

Por tanto, si se quieren crear rutas de forma automática, ¿qué tipo de estrategia se debe seguir? Existen varios tipos de modelos que buscan fundamentalmente minimizar la distancia que tendrá que recorrer el turista para poder realizar la visita, reducir costes, priorizar algunas visitas, etc. (Rodríguez et al., 2012).

Como se ha indicado, una de las bases de conocimiento más importantes sobre el comportamiento de turistas es la que se obtiene a través de estudios de trazabilidad, y ésta será en la que se basa nuestra propuesta de creación de rutas personalizadas. Es decir, se guiará al turista teniendo en cuenta el recorrido habitual que realizan los visitantes de un determinado destino. En nuestra opinión, la utilización de información sobre trazabilidad es un valor añadido que los modelos existentes no consideran.

6.8.2. Utilización de cadenas de Markov para la creación de itinerarios

El sistema debe almacenar información de trazabilidad sobre la ruta que sigue cada turista que visita un determinado destino turístico. En concreto, resulta muy interesante conocer, una vez visitado el ítem i , qué ítem j visita a continuación. Esta información será la que nuestro modelo toma como base para la creación de rutas turísticas.

Esta tarea se puede modelar como un proceso estocástico de tiempo continuo, ya que el estado del tiempo se puede examinar en cualquier momento y no sólo en instantes discretos. Una cadena de Markov (Winston, 2005) es un tipo especial de proceso estocástico de tiempo discreto, donde para $t = 0, 1, 2\dots$ y todos los estados se cumple:

$$\begin{aligned} P(X_{t+1} = i_{t+1} | X_t = i_t, X_{t-1} = i_{t-1}, \dots, X_1 = i_1, X_0 = i_0) &= \\ &= P(X_{t+1} = i_{t+1} | X_t = i_t) \end{aligned}$$

El significado de la ecuación anterior es que la distribución de probabilidad del estado en el tiempo $t+1$ depende de la del estado en el tiempo t y no depende de los estados por los cuales pasó la cadena para llegar al estado i , en el tiempo t . En el estudio de las cadenas de Markov

parte de la hipótesis adicional de que para todos los estados i y j y toda t , $P(X_{t+1} = j|X_t = i) = p_{ij}$, donde p_{ij} es la probabilidad de que el sistema esté en el estado j en el tiempo $t+1$, dado que el sistema está en el estado i en el tiempo t .

La ecuación $P(X_{t+1} = j|X_t = i) = p_{ij}$ indica que la ley de probabilidad que relaciona el estado del siguiente periodo con el estado actual no cambia, o que permanece estacionaria, en el tiempo. Toda cadena de Markov que cumple dicha ecuación se llama cadena estacionaria de Markov. En el contexto de las cadenas de Markov se define q_i como la probabilidad de que la cadena se encuentre en el estado i en el tiempo 0; en otras palabras, $P(X_0 = i) = q_i$. Al vector $q = (q_1, q_2, \dots, q_m)$ se le llama distribución inicial de probabilidad (Winston, 2005).

En el caso que nos ocupa, denotamos como S^1 , el conjunto de ítems que el sistema localiza en el rango de distancia iniciado por el trista y que S^2 denota el subconjunto de ítems pertenecientes a S^1 que cumplen las restricciones de recomendación del turista. $S^2 = \{i_1, i_2, i_3, \dots, i_m\}$ estará formado por un conjunto de ítems que se denominarán estados o ítems visitables por el turista..

Si algún turista ha visitado j después de haber visitado i , diremos que ha ocurrido una transición de i a j . En este caso, se asume que se ha visitado en el tiempo t el punto i y se ha visitado el punto j en el tiempo $t+1$. Se denomina probabilidad de transición (denotada p_{ij}) a la probabilidad de que un turista cualquiera, una vez visita el punto de interés turístico i , el siguiente punto turístico a visitar será el punto j .

Para cada punto de interés turístico i , la serie de valores p_{ij} es una distribución de probabilidad, ya que en cualquier paso puede producirse la visita a alguno de los estados $i_1, i_2, i_3, \dots, i_m$ y son mutuamente excluyentes. Las probabilidades de transición p_{ij} satisfacen:

$$0 \leq p_{ij} \leq 1$$

$$\forall i \in S^2, \sum_{j=1}^m p_{ij} = 1$$

Los diferentes valores p_{ij} son dinámicos, pues se obtienen a partir del historial de visitas de todos los turistas. Obsérvese el ejemplo que se ilustra en la tabla 6.26:

	i_1	i_2	i_3
i_1	0	165	234
i_2	142	0	185
i_3	123	432	0

Tabla 6.26. Trazabilidad de visitas.

Esta tabla registra el orden de las visitas en una zona con tres puntos de interés, de forma que, por ejemplo, el valor 165 de la segunda fila indica que 165 turistas que visitaron i_1 después visitaron i_2 . Se asume, por tanto, que

$$p_{12} = \frac{165}{165 + 234}$$

De esta forma se construye una matriz T con las probabilidades de transición de los estados de S^2 . Se trata de una matriz cuadrada de orden m (siendo m el cardinal de S^2) donde la diagonal principal

contiene valores nulos. Además, para cada fila de la matriz hay una distribución de probabilidad, es decir, la suma de las probabilidades de cada fila es la unidad.

$$T = \begin{pmatrix} 0 & p_{12} & \dots & p_{1m} \\ p_{21} & 0 & \dots & p_{2m} \\ \vdots & \vdots & & \vdots \\ p_{m1} & p_{m2} & \dots & 0 \end{pmatrix}$$

En el ejemplo anterior de la tabla 6.26 tendríamos la siguiente matriz de probabilidad de transición:

$$T = \begin{pmatrix} 0 & 0.41 & 0.59 \\ 0.43 & 0 & 0.57 \\ 0.22 & 0.78 & 0 \end{pmatrix}$$

Estas matrices pueden ser representadas mediante un grafo dirigido que describe las probabilidades de transición. En la figura 6.31 se muestra el grafo correspondiente al ejemplo que estamos utilizando.

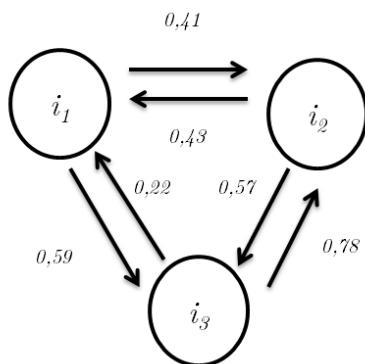


Figura 6.31. Grafo dirigido sobre tres puntos turísticos.

Cada arco tendrá asociado el valor p_{ij} de probabilidad de transición desde el punto de interés turístico i al punto de interés turístico j .

6.8.3. Probabilidad de transición de n etapas. Generación de una ruta de n visitas.

Una medida que puede proporcionar al destino turístico una información muy interesante es la probabilidad de que un turista cualquiera que está visitando el ítem o estado i en el tiempo m , esté tras n visitas en el estado o ítem j . A esta probabilidad se le llama probabilidad en la etapa n de una transición del estado i al estado j (Winston, 2005). Esta información puede servir al destino para saber por ejemplo qué ítems debe promocionar para conseguir un mayor número de visitas. Respecto a la generación de rutas, este valor puede servirnos para poder generar el orden de visita de los n ítems que han sido recomendados a un turista. A continuación se expone c como partiendo de esta medida podemos obtener el orden de visita de los ítems que forman una ruta de n ítems recomendados.

Como se trata de una cadena de Markov estacionaria, esta probabilidad será independiente de m y por tanto:

$$P(X_{m+n} = j | X_m = i) = P(X_n = j | X_0 = i) = p_{ij}^{(n)}$$

donde $p_{ij}^{(n)}$ denota la probabilidad en la etapa n de una transición del estado i al estado j . Por tanto, $p_{ij}^{(0)} = p_{ij}$

Se denomina $p_i^{(0)}$ a la probabilidad de que el turista esté visitando inicialmente el punto i .

Se define $p^{(0)}$ como un vector de probabilidades iniciales con la siguiente estructura:

$$p^{(0)} = (p_1^{(0)}, \dots, p_m^{(0)})$$

Los valores de los elementos del vector $p^{(0)}$ se pueden obtener de diferentes formas:

- a) Caso 1: si la visita inicial está decidida, $p_i^{(0)} = 1$ siendo i el ítem que se está visitando inicialmente, y $p_j^{(0)} = 0.$, para todo ítem j diferente de i . Esta situación sería la idónea cuando el turista tiene claro cuál de los ítems que le han sido recomendados va a visitar en primer lugar.
- b) Caso 2: si la visita inicial no está decidida, $p_i^{(0)}$ será igual al valor de probabilidad de interés de visita del ítem i . Dicho valor se obtiene a partir del valor de preferencia de interés obtenido por el motor de recomendación y debe estar normalizado en una escala entre el 0 y el 1 para poder aplicar las cadenas de Markov.
- c) Caso 3: si la visita inicial no está decidida, se asigna a $p_i^{(0)}$ un valor 1 al ítem que mayor número de visitas ha recibido y un cero al resto. En este caso, el $p_i^{(0)}$ tiene en cuenta las frecuencias de visitas de los turistas que han visitado el destino.

Se define $p_j^{(1)}$ como la probabilidad de visitar el punto j en un solo paso¹¹⁶. Aplicando el teorema de probabilidad total tendríamos:

$$p_j^{(1)} = \sum_{i=1}^m p_i^{(0)} \cdot p_{ij}$$

¹¹⁶ Se asume que no se pueden repetir visitas, por lo que visitado el ítem i , el sistema no podrá recomendar en un paso posterior una nueva visita al mismo ítem i .

Además, para cualquier valor n , $p_{ij}^{(n)}$ se define como la probabilidad de estar en el punto turístico j en la visita n .

De forma análoga a $p^{(0)}$, se define el vector $p^{(1)}$ que almacena la probabilidad de que se alcance cada uno de los ítems después de haber realizado la primera visita:

$$p^{(1)} = (p_1^{(1)}, \dots, p_m^{(1)})$$

$$p^{(1)} = [p_j^{(1)}] = \left[\sum_{i=1}^m p_i^{(0)} \cdot p_{ij} \right] = p^{(0)} \cdot T$$

De la misma manera podríamos expresar:

$$p^{(2)} = p^{(1)} \cdot T = p^{(0)} \cdot T^2$$

$$p^{(3)} = p^{(2)} \cdot T = p^{(1)} \cdot T \cdot T = p^{(0)} \cdot T^3$$

Y, por inducción

$$p^{(n)} = p^{(n-1)} \cdot T = p^{(0)} \cdot T^n$$

Denominamos $p_{ij}^{(n)}$ a la probabilidad de que se visite el ítem j después de n visitas habiendo empezado por el ítem i .

Si se sigue la notación de las cadenas de Markov, se tendría para $n \geq 2$:

$$p_{ij}^{(n)} = P(X_n = j | X_0 = i) = \sum_{k=1}^m P(X_n = j, X_{n-1} = k | X_0 = i)$$

Quedando

$$p_{ij}^{(n)} = \sum_{k=1}^m P(X_n = j, X_{n-1} = k | X_0 = i) =$$

$$\begin{aligned}
 &= \sum_{k=1}^m P(X_n = j | X_{n-1} = k, X_0 = i) \cdot P(X_{n-1} = k | X_0 = i) = \\
 &= \sum_{k=1}^m P(X_n = j | X_{n-1} = k) \cdot P(X_{n-1} = k | X_0 = i) = \\
 &= \sum_{k=1}^m p_{kj}^{(1)} \cdot p_{ik}^{(n-1)} = \sum_{k=1}^m p_{ik}^{(n-1)} \cdot p_{kj}^{(1)}
 \end{aligned}$$

A esta ecuación se le denomina Chapman-Kolmogorov (Winston, 2005).

Por tanto, se verifica que:

$$\left[p_{ij}^{(n)} \right] = T^n$$

Por tanto, el orden de visita de los ítems se encontrará influenciado por las medidas de trazabilidad de las secuencias de visitas de los diferentes ítems. Es importante destacar que solo participan aquellos ítems que han sido recomendados al turista, y que por tanto deben tener un interés adecuado para el mismo.

6.8.4. Potencia enésima de la matriz transición

En la sección anterior se ha descrito una ecuación en la que se necesita calcular la potencia enésima de la matriz transición. En este apartado se describe el mecanismo por el cual se puede obtener algorítmicamente esta matriz minimizando cálculos.

Sea T la matriz de transición de los puntos de interés turísticos recomendados por el sistema pertenecientes a S^2 . Sea I la matriz identidad de la misma dimensión que la matriz T obtenida. Mediante diagonalización de matrices en términos de autovalores y autovectores

se tiene que los autovalores se calculan a partir de la siguiente ecuación (Isaacson y Madsen, 1976; Bhat, 1985):

$$\det(T - \varphi I) = |T - \varphi I| = 0$$

Supóngase que T tiene m autovalores distintos (denominados $\varphi_1, \varphi_2, \dots, \varphi_m$) y que los autovectores correspondientes se denominan r_1, r_2, \dots, r_m , entonces se verifica que (Isaacson y Madsen, 1976; Bhat, 1985):

$$(T - \varphi_i \cdot I)r_i = 0, \quad \text{con } 1 \leq i \leq m$$

Sea C la matriz formada por los autovectores de T , es decir:

$$C = (r_1, r_2, \dots, r_m)$$

$$\begin{aligned} T \cdot C &= T \cdot (r_1, r_2, \dots, r_m) = (T \cdot r_1, T \cdot r_2, \dots, T \cdot r_m) \\ &= (\varphi_1 \cdot r_1, \varphi_2 \cdot r_2, \dots, \varphi_m \cdot r_m) = \\ &= (r_1, r_2, \dots, r_m) \cdot \begin{pmatrix} \varphi_1 & 0 & \dots & 0 \\ 0 & \varphi_2 & \ddots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \varphi_m \end{pmatrix} = C \cdot D \end{aligned}$$

donde D es la matriz diagonal con los autovalores situados en la diagonal principal, pudiéndose calcular T de la siguiente forma:

$$T \cdot C = C \cdot D$$

$$T \cdot C \cdot C^{-1} = C \cdot D \cdot C^{-1}$$

$$T \cdot I = C \cdot D \cdot C^{-1}$$

$$T = C \cdot D \cdot C^{-1}$$

Por tanto, las potencias de T serán:

$$T^2 = (C \cdot D \cdot C^{-1})^2 = C \cdot D^2 \cdot C^{-1}$$

Y, por inducción

$$T^n = C \cdot D^n \cdot C^{-1}$$

donde

$$D^n = \begin{pmatrix} \varphi_1^n & 0 & \dots & 0 \\ 0 & \varphi_2^n & & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & \varphi_m^n \end{pmatrix}$$

Se trata de un método bastante útil a la hora de realizar los diferentes cálculos para obtener la potencia enésima de la matriz de transición (Isaacson y Madsen, 1976; Bhat, 1985).

6.8.5. Trazabilidad turística y cadenas de Markov

En esta sección se presentan algunas medidas aplicables a las cadenas de Markov (probabilidades de estado estables y tiempos medios de primer pasaje) que proporcionan a los destinos turísticos una valiosa información sobre la trazabilidad de los turistas.

6.8.5.1. Clasificación de puntos turísticos en un grafo.

Cuando el número de transiciones a estudiar es grande, las probabilidades de transición de n etapas tienden a estabilizarse (Bhat, 1985). Antes de poder describir con detalle esta afirmación, se aportan algunas definiciones de conceptos propios de los grafos (Winston, 2005), en este caso particularizados para la representación de puntos de interés turístico:

- **Trayectoria:** se denomina trayectoria entre dos ítems i y j , a una sucesión de transiciones que comienza en i y termina en j , de tal

manera que la probabilidad de cada transición de la secuencia es estrictamente mayor que 0.

- Ítem alcanzable: se dice que un ítem j es alcanzable desde un ítem i , si existe al menos una trayectoria de i a j .
- Comunicación entre puntos: se dice que entre un ítem i y un ítem j existe comunicación, si j es alcanzable desde i o i es alcanzable desde j .
- Conjunto cerrado de ítems: se dice que un conjunto S de ítems es conjunto cerrado, si ningún ítem no perteneciente a S es alcanzable desde un ítem de S .
- Camino obligatorio: se denomina camino obligatorio cuando desde un ítem i la probabilidad de transición de ir al ítem j es 1.
- Ítem transitorio: se dice que un ítem i es transitorio, si hay un ítem j alcanzable desde i , pero el ítem i no es alcanzable desde j .
- Ítem recurrente: se dice que un ítem es recurrente si no es transitorio.
- Ítem periódico: un punto i es periódico con periodo $k > 1$, si k es el menor número tal que todas las trayectorias que parten del ítem i y regresan al mismo tienen una longitud múltiplo de k . Si un ítem no es recurrente no es periódico, es decir, es aperiódico.
- Grafo ergódico: si todos los ítems de un grafo son recurrentes, aperiódicos y se comunican entre sí, se dice que es un grafo ergódico.

6.8.5.2. Probabilidad estable aplicada a la trazabilidad.

En esta sección se presenta una medida útil para los destinos turísticos a la hora de realizar un estudio detallado sobre la trazabilidad

de los turistas. Se trata de una medida que permite describir el comportamiento de una cadena de Markov a largo plazo.

Así, el destino puede conocer partiendo de una matriz de transición, cuáles serán las probabilidades de visitas de una serie de puntos dentro de un tiempo. Esta medida es de especial interés, ya que permitirá a los gestores del destino orientar su política turística.

En este apartado se pretende destacar la importancia de la probabilidad estable. Si T es la matriz de transición y el grafo es ergódico, entonces existe un vector de probabilidad estable que se calcula de la siguiente manera:

$$\lim_{n \rightarrow \infty} T^n = \begin{pmatrix} \pi_1 & \pi_2 & \dots & \pi_m \\ \pi_1 & \pi_2 & & \pi_m \\ \vdots & \vdots & & \vdots \\ \pi_1 & \pi_2 & \dots & \pi_m \end{pmatrix}$$

Donde el elemento ij-ésimo de la matriz T^n es $p_{ij}^{(n)}$, cumpliendo que:

$$\lim_{n \rightarrow \infty} T^n = \pi_j$$

Para un valor de n grande, T^n tiende a una matriz con filas iguales (Winston, 2005). Esto quiere decir que después de varias visitas, la cadena de Markov se estabilizará independientemente del estado inicial $p^{(0)}$, existiendo una probabilidad π_j de que nos encontremos en el punto turístico j. Al vector $\pi = (\pi_1, \pi_2, \dots, \pi_m)$ se le llama distribución de estado estable. En el contexto de este trabajo se le denominará distribución de probabilidad de visita estable o también distribución de

equilibrio estable. Así, para un valor grande de n y para cualquier ítem o punto de interés i :

$$p_{ij}^{(n+1)} \cong p_{ij}^{(n)} \cong \pi_j$$

Como se indicó en la sección 6.8.3 $p_{ij}^{(n+1)}$ hace referencia al producto de la fila i de la matriz T^n con la columna j de T, y se calcula de la siguiente forma:

$$p_{ij}^{(n+1)} = \sum_{k=1}^m p_{ik}^{(n)} \cdot p_{kj}$$

Si el valor de n es grande entonces (Isaacson y Madsen, 1976):

$$\pi_j = \sum_{k=1}^m \pi_k \cdot p_{kj}$$

En forma matricial, podemos escribir la ecuación anterior de la siguiente manera:

$$\pi = \pi \cdot T$$

El principal inconveniente de la resolución de este sistema de ecuación lineal es que, aplicando el teorema de *Rouché-Fröbenius*, tendríamos que el rango de la matriz de coeficientes y el de la matriz ampliada será inferior al número de incógnitas, con lo que el sistema tendría infinitas soluciones, ya que el rango de la matriz T será siempre menor o igual a m-1, por lo que para poder obtener unos valores únicos de probabilidades estable debe cumplirse que (Winston, 2005):

$$p_{i1}^{(n)} + p_{i2}^{(n)} + \cdots + p_{im}^{(n)} = 1$$

Al tender n a infinito tendremos que:

$$\pi_1 + \pi_2 + \cdots + \pi_m = 1$$

A continuación se muestra la aplicación de estos resultados al ejemplo de la sección 6.8.2. Recordamos que la matriz T tiene los siguientes valores:

$$T = \begin{pmatrix} 0 & 0.41 & 0.59 \\ 0.43 & 0 & 0.57 \\ 0.22 & 0.78 & 0 \end{pmatrix}$$

Por lo que resolviendo el sistema:

$$(\pi_1 \quad \pi_2 \quad \pi_3) = (\pi_1 \quad \pi_2 \quad \pi_3) \cdot \begin{pmatrix} 0 & 0.41 & 0.59 \\ 0.43 & 0 & 0.57 \\ 0.22 & 0.78 & 0 \end{pmatrix}$$

Se tiene que:

$$\begin{aligned} \pi_1 &= \pi_1 \cdot 0 + \pi_2 \cdot 0.43 + \pi_3 \cdot 0.22 \\ \pi_2 &= \pi_1 \cdot 0.41 + \pi_2 \cdot 0 + \pi_3 \cdot 0.78 \\ \pi_3 &= \pi_1 \cdot 0.59 + \pi_2 \cdot 0.57 + \pi_3 \cdot 0 \\ \pi_1 + \pi_2 + \pi_3 &= 1 \end{aligned}$$

Obtenido como vector de probabilidad estable
 $\pi = (0.24692, 0.38688, 0.3662)$

El ejemplo propuesto es pequeño en cuanto al número de elementos, pero es significativo para el estudio propuesto. Una pregunta que podemos hacernos es en qué momento se obtienen en las cadenas de Markov la probabilidad estable. Como regla general, puede afirmarse que cuando la matriz T tiene pocos elementos que queden cerca de 0 o de 1, habitualmente se alcanzará de forma rápida. Al comportamiento de una cadena de Markov antes de alcanzar el estado estable se llama comportamiento transitorio. Sin embargo, es bueno conocer el estado

estable para saber con exactitud la probabilidad de encontrarse en un estado determinado (Winston, 2005).

6.8.5.3. Visitas promedio de primer pasaje

Otra medida que puede resultar interesante al destino es conocer el número de ítems medio que se visita hasta llegar a visitar un determinado ítem o punto de interés estando en otro ítem determinado. Para ello se define en un grafo ergódico, el valor m_{ij} al número esperado de visitas antes de visitar el ítem j teniendo en cuenta que se está visitando el ítem i . A este valor m_{ij} , formalmente se le denominará número promedio de primer pasaje desde el ítem i hasta el ítem j (Isaacson y Madsen, 1976).

Si se supone que el turista está en el ítem i , entonces existe una probabilidad p_{ij} de visitar directamente el ítem j después de haber visitado el ítem i .

Supóngase que no se visita j desde i de forma directa. Existe entonces k , tal que $k \neq j$, de forma que se visitará k desde i con una probabilidad p_{ik} . En este caso se necesitará un promedio de $1 + m_{kj}$ visitas, para poder pasar efectivamente del ítem i al ítem j . De esta manera tenemos (Isaacson y Madsen, 1976):

$$m_{ij} = p_{ij}^{(1)} + \sum_{k \neq j} p_{ik} \cdot (1 + m_{kj})$$

Como se sabe que:

$$p_{ij} + \sum_{k \neq j} p_{ik} = 1$$

Se tiene entonces que:

$$m_{ij} = 1 + \sum_{k \neq j} p_{ik} \cdot m_{kj}$$

Con lo cual tenemos un nuevo sistema de ecuaciones lineal donde podemos encontrar los tiempos promedios de primer pasaje, obteniendo como solución que (Isaacson y Madsen, 1976):

$$m_{ii} = \frac{1}{\pi_i}$$

Así, por ejemplo si m_{ij} valiera 3 significaría que, habiendo visitado un turista el ítem i , visitará como promedio otros tres ítems antes que el ítem j .

6.8.6. Algoritmo para la creación de rutas turísticas

En este apartado se presenta el algoritmo mediante el cual el sistema propuesto crea una ruta a partir de un conjunto de ítems recomendados por alguno de los motores de recomendación presentados en las secciones anteriores.

El proceso de creación de rutas consta de los siguientes pasos:

Paso 1: Selección de datos para generar ruta. El turista debe decidir si quiere empezar la visita por algún ítem en concreto o si quiere que se le recomiende un ítem de inicio de acuerdo a las frecuencias de visitas de los turistas que han visitado el destino o según los valores de predicción de interés que se han obtenido de los motores de recomendación. Esta elección determinará los valores del vector $p^{(0)}$.

Paso 2: Generación del orden de los ítems en la ruta. A continuación, se realiza el cálculo de $p^{(1)}$, eligiendo como segundo ítem de la ruta aquel elemento del vector que tenga el mayor valor y que no haya sido visitado anteriormente. De forma análoga se eligen los siguientes ítems de la ruta calculando los vectores $p^{(2)}, \dots p^{(n-1)}$.

A continuación, se muestra sobre el ejemplo propuesto la influencia que tiene la forma de definir el vector de probabilidades iniciales $p^{(0)}$. Así, si partimos de la siguiente matriz de transición:

$$T = \begin{pmatrix} 0 & 0.41 & 0.59 \\ 0.43 & 0 & 0.57 \\ 0.22 & 0.78 & 0 \end{pmatrix}$$

Caso 1: Supóngase que el turista ha decidido visitar en primer lugar o le ha sido recomendado preferentemente iniciar su visita en el ítem 1, en este caso si la ruta que solicita está compuesta por tres ítems, se tendría que:

$$p^{(0)} = (1 \quad 0 \quad 0)$$

Entonces, los valores de $p^{(1)}$ y $p^{(2)}$ son

$$\begin{aligned} p^{(1)} &= (1 \quad 0 \quad 0) \cdot \begin{pmatrix} 0 & 0.41 & 0.59 \\ 0.43 & 0 & 0.57 \\ 0.22 & 0.78 & 0 \end{pmatrix} = (0 \quad 0.41 \quad 0.59) \\ p^{(2)} &= (0 \quad 0.41 \quad 0.59) \cdot \begin{pmatrix} 0 & 0.41 & 0.59 \\ 0.43 & 0 & 0.57 \\ 0.22 & 0.78 & 0 \end{pmatrix} \\ &= (0.3061 \quad 0.4602 \quad 0.2337) \end{aligned}$$

Este cálculo informa de que después de visitar el ítem i_1 , se debería visitar i_3 y luego i_2 .

Caso 2: En cambio, si se toman como punto de partida los valores de probabilidad de interés teniendo en cuenta las visitas realizadas por el turista en su historial (véase el ejemplo de la sección 6.6.4, del cual se han tomado los datos), se tendría que:

$$p^{(0)} = (0.0252 \quad 0.1272 \quad 0.0364)$$

Con lo que

$$\begin{aligned} p^{(1)} &= (0.0252 \quad 0.1272 \quad 0.0364) \cdot \begin{pmatrix} 0 & 0.41 & 0.59 \\ 0.43 & 0 & 0.57 \\ 0.22 & 0.78 & 0 \end{pmatrix} = \\ &= (0.062704 \quad 0.038724 \quad 0.087372) \end{aligned}$$

Este cálculo nos informa de que si se solicitase una ruta sugerida formada por dos puntos turísticos, la elección sería visitar en primer lugar el ítem i_2 , dado que tiene un mayor nivel de preferencia según las evaluaciones realizadas por el turista y la siguiente visita debería realizarse al ítem i_3 , dado que en la primera transición es el valor mayor (0.087372).

Caso 3: En este caso el proceso es idéntico al indicado en el caso 1, con la única diferencia que se le asigna el valor 1 al ítem que mayor número de visitas haya tenido, y asignando un 0 al resto.

Paso 3: Planificación de la ruta generada personalizada al turista. Una vez que se le muestra al turista la ruta sugerida, se le debe solicitar información sobre duración de la visita, coste económico que está dispuesto a asumir, etc., para de esta forma se pueda generar la ruta definitiva teniendo en cuenta esas consideraciones. El motivo de solicitar a posteriori esta información tiene como objetivo que haga una

estimación más adecuada una vez que tenga una propuesta de ítems a visitar. Por ejemplo, quizá un turista tenga pensado dedicar un tiempo a la visita, pero al comprobar los ítems quizá cambie su opinión inicial.

Por tanto, una vez generada la lista de puntos que forman la ruta, hay que comprobar si los ítems seleccionados cumplen los criterios del turista. Por ejemplo, si dispone de ocho horas para realizar la ruta y no quiere gastar más de 30 euros por persona, ¿la ruta generada es válida? Quizá, el sistema recomiende visitar cinco puntos de interés y los dos primeros ya tengan un coste de 40 euros, lo que haría imposible completar la ruta.

Sin pérdida de generalidad, en adelante asumiremos los siguientes criterios de validación de rutas:

- Distancia: es importante tener en cuenta la distancia existente entre dos ítems. Cuando las distancias requieren el uso de un medio de transporte, disminuirá el tiempo y aumentará el coste económico. De este criterio se derivan dos objetivos: minimizar el coste y minimizar el tiempo de transporte.
- Coste: la visita de una determinado ítem o actividad turística puede tener un coste. En caso de que sea así, en la mayoría de los casos éste dependerá del número de personas que realizan la visita. De este criterio, obviamente, se deriva el objetivo de minimizar el coste.
- Duración: la visita de un ítem tiene una duración. Se utilizará como valor la estimación media que tiene la visita de ese ítem. De este criterio se deriva el objetivo de maximizar la satisfacción

del turista. Una visita corta por tener que visitar otros ítems puede frustrar al turista.

- Preferencias: las visitas tienen que realizarse teniendo en cuenta los gustos y preferencias del turista. De este criterio se deriva igualmente el objetivo de maximizar las preferencias del turista.

En consecuencia, se proponen los siguientes atributos para la generación de la ruta diaria son:

- a) Tiempo de duración de la visita.
- b) Intervalo de tiempo para comer o avituallamiento.
- c) Presupuesto para realizar la visita.

Una vez que se obtiene la lista ordenada de ítems que componen la ruta (Paso 2), es necesario comprobar si ésta satisface los criterios anteriores. Debe mencionarse que se asume que todos los ítems sugeridos por el sistema cumplen los atributos contextuales, como de hecho se ha tenido en cuenta en la definición de todos los motores de recomendación de este trabajo. Dado que los ítems que forman parte de la ruta han sido seleccionados por nuestro motor de recomendación, la conformidad con el contexto está garantizada. No obstante existe una excepción, que viene dada por el horario de visita que se asigna al ítem en la ruta. Debe comprobarse que sea compatible con el horario de apertura del ítem.

Sea L la lista ordenada de n ítems sugeridos al turista como parte de la ruta generada en el paso anterior.

$$L = \{i_1, i_2, \dots, i_n\}$$

El algoritmo para la validación de ruta y generación de la ruta personalizada consiste en la aplicación de los siguientes pasos:

Paso 3.1: Eliminar aquellos ítems cuyo coste supere al total indicado por el turista.

Paso 3.2: Estudiar para cada ítem qué duración y coste tendría ir hasta el siguiente ítem de la lista, es decir, qué duración y coste tiene esta transición. Para ello se incluyen en la lista de ítems ficticios de transición, que tienen un coste y duración. A esta lista la denominaremos L_1

$$L_1 = \{i_1, t_1, i_2, t_2, \dots, i_n\}$$

donde i son las actividades y t las transiciones entre esas actividades.

Paso 3.3: Si la suma total de la duración de los ítems de la lista L_1 supera el máximo indicado por el turista, se eliminarán los últimos ítems de la lista, hasta que la duración total no supere el tiempo estimado por el turista. A la lista de ítems resultante (una sublistas de L_1) la denominaremos L_2 . Mientras, los ítems eliminados forman otra sublistas denominada L_E .

Paso 3.4: Estudiar si algunos de los ítems incluidos en la lista L_2 sería incompatible¹¹⁷ con el horario asignado en la ruta. En estos casos el ítem se intercambiará con aquél ítem o punto más cercano en la lista L_2 que

¹¹⁷ Nuestros motores de recomendación generan recomendaciones compatibles contextualmente con diferentes atributos contextuales. Uno de ellos es el atributo contextual relacionado con la hora a la que se realiza la visita. Quizá un ítem que ahora es realizable dentro de unas pocas horas no puede realizarse. Esta comprobación contextual es necesaria. Diremos entonces que un ítem será incompatible con el horario si por ejemplo el horario que se le ha asignado en la ruta no se encuentra abierto.

permite su compatibilidad. En caso de incompatibilidad total, se eliminará de la lista, sustituyéndola por el primer ítem de la lista L_E . Este ítem será ahora el último de la lista L_2 .

Paso 3.5: Si la suma total del coste supera el máximo indicado por el turista, habrá que eliminar la transición y el ítem último por el primer ítem eliminado de la lista L_2 , pero no se añade a la lista L_E . Ahora se vuelven a repetir los pasos volviendo a repetir el paso 3.3 al 3.5 hasta que la lista cumpla los requisitos de tiempo y de coste asumiendo que se asigna a L_1 la lista L_2 y la lista L_E sigue manteniendo sus ítems en el resto de pasos del algoritmo.

6.9. Consideraciones finales.

En la mayoría de los motores de recomendación propuestos se han utilizado los mismos ítems en los diferentes ejemplos que ilustran el funcionamiento de los diferentes motores de recomendación (excepto el presentado en la sección 6.6 que para evitar cálculos engorrosos lo hemos reducido a un número inferior de ítems). Una conclusión muy importante que se puede extraer es que se hace necesaria la utilización de métodos híbridos, pues los resultados difieren de forma clara de un modelo a otro.

En la siguiente tabla de resumen se muestran los cinco ítems recomendados en orden descendente en los diferentes motores:

Motor de recomendación	Ítems recomendados
Basado en contenido (sección 6.3)	$i_{10}, i_6, i_2, i_4, i_3$
Demográfico (sección 6.4)	$i_{10}, i_2, i_8, i_1, i_5$
Colaborativo (sección 6.5)	$i_5, i_{10}, i_2, i_6, i_1$
Grupal (sección 6.7)	$i_{10}, i_6, i_2, i_1, i_3$

Tabla 6.27. Resumen de ítems recomendados

En la figura 6.32 se puede observar gráficamente el valor de predicción de interés recomendado a cada uno de los ítems dependiendo del motor de recomendación utilizado. La altura de la columna es proporcional al valor de predicción de interés del ítem. Como puede observarse los ítems i_7 e i_9 no obtienen ningún resultado positivo. El motivo es que no cumplen adecuadamente los diferentes parámetros contextuales requeridos.

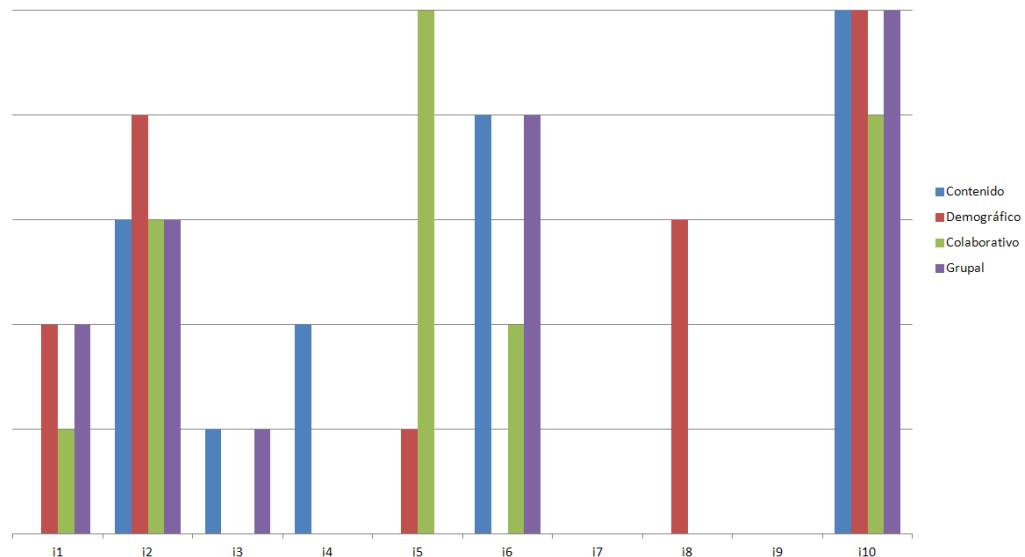


Figura 6.32. Predicción de interés de los ítems en cada motor de recomendación

Debido a que los motores utilizan información diferente, los resultados también difieren. Por tanto, la utilización de métodos híbridos configurables por parte del mismo turista es una solución adecuada para conseguir que las recomendaciones generadas al turista maximicen las preferencias del mismo.

Conclusiones y trabajos futuros

A continuación se exponen las principales conclusiones extraídas de este trabajo de investigación. Posteriormente se presentan las líneas de investigación y trabajo futuras que se plantean.

Conclusiones

El turista del siglo XXI demanda cada vez más información adaptada a sus preferencias, por lo que las empresas y las organizaciones públicas de gestión del turismo deben proporcionar herramientas para prestar un mejor servicio y ofrecer aquellos productos que son más afines al turista.

Además, dichas herramientas pierden gran parte de su potencial si funcionan de forma aislada. Es por ello que existe una necesidad de colaboración entre los diferentes agentes del sector turístico (empresas, organismos públicos, centros de investigación, etc.) para crear arquitecturas que permitan la integración de los diferentes sistemas de información turística.

La hipótesis realizada en la presente tesis doctoral relativa a la implantación de herramientas innovadoras en los destinos turísticos, facilitará el acceso a la información y por tanto generará ventajas competitivas en los destinos que las implanten. Se ha desarrollado el modelo RAMCAT, un ambicioso marco de trabajo para sistemas de información turística, que combina de forma óptima las funcionalidades de recomendación y realidad aumentada. En particular, se ha mostrado la importancia de incorporar recomendaciones personalizadas basadas en el contexto, perfeccionando éstas con el conocimiento extraído tanto de las interacciones del turista con el sistema, como de las valoraciones que éste haga de dichas recomendaciones.

El trabajo realizado ha permitido mostrar la importancia de la incorporación de sistemas de recomendación en las herramientas de realidad aumentada diseñadas para el turismo, ya que permiten ofrecer información personalizada respecto a las preferencias del turista. De esta forma, los turistas podrán encontrar de forma fácil y rápida aquellos puntos de interés que desean visitar o utilizar, sin verse desbordados por una sobreabundancia de información.

Es muy habitual que los sistemas de recomendación generen sus predicciones a partir de información de las visitas realizadas previamente por el turista que solicita la recomendación. Cuando un turista visita un destino por primera vez no se dispone de información sobre anteriores visitas (problema conocido como *cold start*), por lo que es necesario incorporar y combinar técnicas de recomendación que permitan realizar recomendaciones tanto si el sistema tiene información

sobre visitas anteriores como si no. En este trabajo se han diseñado e incorporado al marco de trabajo motores de recomendación basados en técnicas demográficas y basadas en contenido que permiten generar recomendaciones sin información de visitas anteriores del turista, lo cual soluciona el problema *cold start* y supone una mejora sobre otros modelos de recomendación.

Actualmente, la mayoría de los motores de recomendación no tienen en cuenta información del contexto a la hora de generar las predicciones. Una aportación esencial de la propuesta presentada es la utilización de atributos contextuales, los cuales permiten obtener mejores recomendaciones. Estos atributos son utilizados en dos fases:

- En primer lugar, se ha diseñado una técnica de pre-filtrado contextual que conseguirá reducir drásticamente el número de puntos de interés que constituyen el conjunto de entrada al sistema de recomendación, mejorando así la eficiencia computacional al reducir el número de operaciones requeridas. Esta técnica de pre-filtrado contextual basado en implicaciones es común a todos los motores de recomendación propuestos, y utiliza técnicas difusas y análisis formal de conceptos. Estas implicaciones relacionan atributos contextuales con características de puntos de interés.

- En segundo lugar, una vez realizada la etapa de pre-filtrado, los atributos contextuales vuelven a ser utilizados por todos los motores de recomendación propuestos para garantizar que los puntos de interés seleccionados sean compatibles con el contexto del turista,

Un aspecto fundamental para generar recomendaciones precisas es disponer de información adecuada sobre las valoraciones que hacen los turistas. Son muchos los motores de recomendación que utilizan escalas exclusivamente numéricas para que el turista y los expertos puedan valorar un determinado ítem o una característica. En muchos casos resulta difícil modelar esas valoraciones en forma de información cuantitativa y se hace necesario gestionar valoraciones subjetivas.

Para resolver este problema, se propone la utilización de información lingüística, técnica muy apropiada para modelar la información subjetiva propia tanto de las valoraciones como de las preferencias del turista, que con frecuencia tienen asociadas un alto grado de incertidumbre. El uso de diferentes escalas lingüísticas es una solución muy adecuada para poder valorar diferentes características y propiedades de un determinado punto de interés.

En la mayoría de los modelos existentes no se tiene en cuenta el problema de la generación de recomendaciones para grupos turísticos, siendo el viaje en grupo la opción más habitual cuando se visita un destino. Nuestra propuesta soluciona este inconveniente, permitiendo generar recomendaciones adecuadas y compatibles con los diferentes miembros de un grupo, además de permitir la generación de rutas tanto grupales como individuales.

Otro aspecto relevante para las organizaciones de gestión de destinos es la recopilación de información sobre los turistas. El modelo propuesto, al tiempo que genera las recomendaciones, construye una base de conocimiento turístico, a partir de la cual se puede conocer de

manera precisa la realidad de un determinado punto de interés y obtener información sobre trazabilidad turística.

Todo el trabajo realizado a lo largo de esta tesis doctoral permite afirmar que la aplicación en un destino turístico del marco de trabajo propuesto proporcionará beneficios como:

- a) Ofrecer a los turistas una nueva e innovadora herramienta para personalizar su viaje, incluyendo las actividades de preparación, planificación y visita al destino.
- b) Facilitar la identificación de los recursos turísticos del destino.
- c) Aportar a las instituciones y organismos del destino conocimiento sobre las preferencias de los turistas, trazabilidad y gestión del flujo turístico.
- d) Obtener un conocimiento eficiente de la realidad del destino, permitiendo orientar la política turística, promoviendo la localización y el acceso a la información de los recursos turísticos.

Por último, hay que realizar una reflexión muy importante sobre el desarrollo de nuevos sistemas de información turísticos, en concreto afirmamos que dichos sistemas no aportarán todas sus ventajas potenciales si no consiguen integrarse con los sistemas ya existentes. Nuestra opinión es que sólo la interoperabilidad de los diferentes sistemas de un destino turístico permite que la gestión de la información sea sostenible y que tenga un nivel de calidad satisfactorio. Todo esto, redundará en beneficio de las empresas, destinos y usuarios.

Líneas de investigación y trabajo futuro

En el ámbito de la transferencia de resultados de investigación, se plantea la importante tarea futura de la integración del sistema propuesto en la plataforma tecnológica de la Costa del Sol Occidental¹¹⁸, desarrollada por nuestro grupo de investigación SICUMA¹¹⁹. Esta implantación conllevará importantes beneficios para los turistas, ya que incorporará a dicha plataforma las funcionalidades de recomendación de puntos de interés y de generación de rutas adaptadas a las preferencias del turista. El destino Costa del Sol Occidental poseerá una innovadora herramienta que contribuirá a la mejora de la competitividad del destino.

Por otra parte, en lo que respecta a la continuidad de las líneas de investigación de este trabajo, en un futuro cercano pretendemos desarrollar un módulo de recomendación basado en redes sociales. Este motor generará las recomendaciones teniendo en cuenta la información extraída mediante el rastreo de las redes sociales, aplicando herramientas de procesamiento de lenguaje natural y análisis de sentimientos para identificar las preferencias (positivas y negativas) del turista y de sus amigos.

Para finalizar queremos resaltar que el trabajo desarrollado en esta tesis también abre líneas de investigación para mejorar los modelos presentados, en concreto nos proponemos:

¹¹⁸ www.costadelsoloccidental.org

¹¹⁹ www.sicuma.uma.es

- a) Estudiar nuevas técnicas de hibridación que permita optimizar los resultados de las recomendaciones y que permitan mejorar los tiempos de ejecución para lograr una mayor eficiencia en la generación de resultados.
- b) Dotar de mayor simplicidad a las estructuras de preferencias de los turistas. Para ello se hará necesario mejorar los procesos de adquisición de información.
- c) Profundizar en aspectos relacionados con los procesos de captura de datos relativos a la personalidad de los turistas, lo que permitirá mejorar las recomendaciones grupales.
- d) Profundizar en el aprovechamiento de la base de conocimiento del turista generada en el sistema propuesto, para generar un sistema optimizado de gestión de las relaciones con los turistas y contribuir en el desarrollo de un verdadero *smart destination*.
- e) Incorporar técnicas basadas en implicaciones difusas en los motores de recomendación, lo que permitirá dotarlos de capacidad de razonamiento, de manera que permitirán al turista conocer los motivos por los que se le ha recomendado una serie de puntos determinados.

Anexo I

Ontología RAMCAT

En el presente anexo se va a incluir el glosario de entidades que la componen, representada mediante la taxonomía o jerarquía de dichas entidades.

I.1. Taxonomía de entidades

- OWL: Thing
 - SistemaRecomendación
 - TipoDestino
 - DestinoPlaya
 - DestinoFamiliar
 - DestinoLagos
 - DestinoCiudad
 - DestinoMontaña
 - DestinoRural
 - DestinoRío

- DestinoVerano
- DestinoInvierno
- DestinoCultural
- Entorno
 - Interior
 - Exterior
 - Mezcla
- RegiónGeográfica
 - País
 - Región
 - Provincia
 - DestinoTurístico
 - Ciudad
 - Localidad
 - Pedanía
 - CentroHistórico
 - Afueras
- POI
 - Alojamiento
 - Apartamento
 - Camping
 - Chalet
 - CasaInvitados
 - Hostal
 - Hotel
 - HotelLujo

- Hotel5E
- Hotel4E
- Hotel3E
- Hotel2E
- Hotel1E
- Residencia
- Resort
- Monumentos y Patrimonio
 - EdificioHistórico
 - Castillo
 - Palacio
 - Alcazaba
 - Templo
 - OtroEdificioHistórico
 - MonumentoCivil
 - Biblioteca
 - Ayuntamiento
 - Auditorio
 - OtroMonumentoCivil
 - MonumentoIndustrial
 - Bodega
 - Depósito
 - Fábrica
 - Lonja
 - Oficina
 - Puente

- OtroMonumentoIndustrial
- LugaresPintorescos
 - Barrios
 - Calles
 - Rincones
 - Caminos
 - Miradores
 - OtrosLugaresPintorescos
- MonumentoReligioso
 - Catedral
 - Iglesia
 - Monasterio
 - Mezquita
 - Ermita
 - CasaHermandad
 - Capilla
 - Cementerio
 - OtroEdificioReligioso
- PatrimonioArqueología
 - CiudadAntigua
 - VillaAntigua
 - YacimientoArqueológico
 - Torre
 - Cueva
 - Muralla
 - Necrópolis

- RestosArqueológicos
- Castillos
- Ruina
- Termas
- Basílicas
- Fuerte
- MuseoArqueológico
- OtrosPatrimonio
 - Escultura
 - Busto
 - Fuente
 - Plaza
 - Mural
- Ecología
 - Parques y plazas
 - Parque
 - Plaza
 - Fuente
 - Jardín
 - RecintoInfantil
 - OtroParquePlaza
 - Zoológicos
 - Abiertos
 - Cerrado
 - Mixto
 - Playas

- PlayaAislada
- Cala
- PlayaUrbana
- PlayaSuburbana
- PlayaNudista
- OtrasPlayas
- TurismoActivo
 - Alpinismo
 - Submarinismo
 - Senderismo
 - Otros
- AtraccionesNaturales
 - Cueva
 - Jardín
 - ParqueNatural
 - Río
 - Lago
 - Montaña
 - Miradores
 - Ruta
 - Caminos
 - OtrosAtraccionesNaturales
- ActividadesDeportivas
 - Recintos
 - CampoFutbol
 - Polideportivo

- PalacioDeportes
- CampoGolf
- Pabellón
- PlazaToros
- CiudadDeportiva
- PistasDeportivas
- Hipódromos
- OtrosRecintos

- Deportes

- CampoAtletismo
- CampoFútbol
- ClubHielo
- ClubDeportivo
- ClubNáutico
- ClubTenis
- ClubSubmarinismo
- ClubAlpinismo
- ClubSenderismo
- ClubHípico
- ClubPádel
- ClubGolf
- Gimnasio
- Hipódromo
- Polideportivo
- Rocódromo
- OtrosDeportes

- Ocio
 - Actividades
 - Arqueológica
 - Enoturismo
 - VisitaMuseo
 - Fiesta
 - Compras
 - Panorámica
 - Rafting
 - Rappel
 - ActividadesCulturales
 - Museos
 - MuseoGeneral
 - Escultura
 - Pintura
 - OtrosMuseos
 - SalasExposiciones
 - Colección
 - CasaNatal
 - CentroCultural
 - CasaCultura
 - SalaCultural
 - Tradiciones
 - OtrasActividadesCulturales
 - Deportes
 - Playa

- Compras
 - Tiendas
 - CentroComercial
 - Mercadillo
 - Mercado
 - OtrosCompras
- Entretenimiento
 - Aquarium
 - ParqueAcuático
 - Teatro
 - Zoológico
 - ParqueTemático
 - ParqueAtracciones
 - Feria
 - OtroEntretenimiento
- Juegos
 - Casino
 - Bingo
 - Hipódromo
 - OtrosJuegos
- VidaNocturna
 - Bar
 - Club
 - Pub
 - Disco
 - Karaoke

- SalaBaile
- OtrosVidaNocturna
- AcontecimientosTuristicos
 - Eventos
 - Festival
 - FestivalCine
 - FestivalMúsica
 - FestivalTeatro
 - Musical
 - Cine
 - Teatro
 - SalaConciertos
 - Auditorios
 - Fútbol
 - Golf
 - Baloncesto
 - Balonmano
 - Natación
 - FestivalTaurino
 - OtrosEventos
 - InstalacionesTurísticas
 - Gastronomía
 - Bar
 - Cafetería
 - Pastelería
 - Tetería

- Heladería
- Restaurante
 - Pizzería
 - Marisquería
 - Chiringuito
 - Mesón
 - Venta
 - Freiduría
 - Autor
 - Argentino
 - Asiático
 - Europeo
 - Americano
 - Burger
 - ComidaRápida
 - Mexicano
 - OtroRestaurante
- Relajación
 - Baños
 - Spa
 - Masaje
 - SaludyBelleza
 - OtrosRelajación
- Puertos
- ServiciosTurísticos
 - Banco

- AlquilerCoches
 - CajerosAutomáticos
 - Hospital
 - Gasolineras
 - Policía
 - Correos
 - CentrosComerciales
 - OficinaTurismo
 - OtrosServiciosTuristicos
- Transporte
 - Aeropuerto
 - EstaciónAutobuses
 - EstaciónMetro
 - EstaciónTaxi
 - EstaciónTren
 - ParadaAutobús
 - OtrosTransporte
 - Puertos
 - PuertoDeportivo
 - PuertoPesquero
 - OtroPuertos
 - Rutas
 - Teleférico
 - Caballos
 - Tren
 - Burro

- OtrasRutas
 - Economía
 - Compras
 - Banco
 - Cajeros
- RecursoInternet
 - RecursosArchivos
 - PDF
 - DOCX
 - TXT
 - 3D
 - RecursosWeb
 - PaginaWeb
 - PWBlog
 - PWFoto
 - PWContenido
 - PWWiki
 - ContenidoPáginaWeb
 - Multimedia
 - Audio
 - Imagen
 - Video
 - Animación
 - Recreación
 - Comentarios
 - Críticas

- Foros
- SitioWeb
 - Blog
 - SitioOficial
 - SitioFotos
 - SitioForos
 - RedesSociales
 - Facebook
 - Twitter
 - SitioVideos
 - Sitio3D
- Idioma
 - Inglés
 - Francés
 - Alemán
 - Árabe
 - Chino
 - Español
 - OtroIdioma
- Usuarios
 - Turista
 - Grupo
 - MiembroGrupo
 - PropietarioGrupo
- Reservas
 - ReservaAlojamiento

- ReservaRestaurante
- Ticket
- ReservaTransporte
- DescriptorTurístico
 - Negocios
 - Familiar
 - Cultural
 - Joven
 - Aventura
 - Naturaleza
 - Solitario
 - Ocio
 - Descanso
 - OtroDescriptor
- AtributoDemográfico
 - NivelCultural
 - Edad
 - NivelEconómico
 - Estado
- Contexto
 - Fecha
 - Hora
 - Temperatura
 - Clima
 - Compañía
 - Niños

- Idiomas
- Minusvalía
- Caluroso
- Lluvia
- OtrosContexto

Anexo II

Prototipo de herramienta basada en el modelo RAMCAT

II.1. Introducción

En el presente anexo se describen las características básicas del prototipo basado en el modelo RAMCAT desarrollado para su aplicación al destino turístico de la Costa del Sol. Actualmente se encuentran implementados de forma independiente los diferentes motores de pre-filtrado y de recomendación, estando en fase de estudio el diseño de la herramienta de realidad aumentada que utilizará dichos motores.

El sistema propuesto tiene como principales características la utilización de atributos contextuales y la aplicación de diferentes motores de recomendación. La base teórica de los distintos modelos de

pre-filtrado y recomendación se ha detallado en el capítulo seis y en éste anexo vamos a describir cómo los utilizarían en distintas situaciones los turistas para solicitar una recomendación.

La mayoría de los sistemas de recomendación turísticos presentan una serie de inconvenientes que han sido solucionados en el modelo propuesto de la siguiente forma:

- a) Muchos de los sistemas de recomendación requieren que el usuario haya valorado diferentes productos con anterioridad a la solicitud de recomendación. El método propuesto soluciona de forma efectiva el problema de arranque en frío (ya citado con anterioridad) que afecta a muchos métodos clásicos de recomendación. La solución viene dada por el hecho de que varios de los motores de recomendación incluidos en nuestro marco de trabajo no requieren que el usuario haya calificado previamente otros productos.
- b) La mayoría de los sistemas de recomendación no utilizan atributos contextuales para la generación de recomendaciones. El modelo propuesto permite realizar un pre-filtrado contextual que reduce el número de ítems con los que trabajan los motores de recomendación, lo que mejorará los tiempos de respuesta. Los valores de los atributos contextuales pueden ser obtenidos mediante interacciones con otros sistemas (posición y clima por ejemplo) o mediante la introducción explícita de dicha información por parte del turista (compañía y características del grupo por ejemplo).

- c) Muchos de los sistemas de recomendación sólo gestionan información numérica de los puntos de interés. Nuestro modelo utiliza diferentes escalas (numéricas y lingüísticas) para describir de forma clara las características de un producto, lo que permite que los usuarios del sistema puedan describir de forma precisa sus expectativas y expresar de forma precisa sus valoraciones.
- d) Es habitual que la mayoría de los sistemas de recomendación no tengan en cuenta las preferencias de los acompañantes de un usuario del sistema. En el modelo propuesto, se permite la generación de recomendaciones para grupos, teniendo en cuenta los atributos contextuales del mismo.
- e) La utilización de información histórica es muy relevante en muchos modelos de recomendación. Siendo una información muy válida y útil, puede ser en determinadas situaciones un elemento débil del modelo, ya que reduce la capacidad de sorpresa al usuario. Nuestra propuesta permite que el turista pueda solicitar recomendaciones que difieran en cierta medida de sus preferencias habituales, consiguiendo de esta forma que la información histórica no sea relevante y no sea utilizada en la generación de recomendaciones. El módulo de recomendación basado en contenido sin memoria posibilitará esta característica.
- f) La utilización de la información demográfica de los turistas, aunque muy útil, hace que la mayoría de las veces las recomendaciones sean siempre muy similares. Nuestra propuesta permite introducir un factor de sorpresa en las recomendaciones demográficas con objeto de permitir que el turista pueda visitar

otro tipo de puntos bien valorados por otros usuarios. La ventaja de nuestro modelo es que en la recomendación que se ofrece solo algunos de los ítems recomendados aportan sorpresa, mientras que el resto se ajustará a las características demográficas del usuario.

En este anexo se presenta el funcionamiento básico del prototipo del sistema. En la primera sección se describe el esquema de recomendación del sistema, mientras que en la segunda sección se detallará el funcionamiento de la interfaz de usuario del prototipo, acabando con las conclusiones.

II.2 Módulos de recomendación y funcionamiento básico

Dado que el sistema propuesto utiliza diferentes motores de recomendación, es necesaria la aplicación de una técnica de hibridación. Como ya se indicó, las técnicas de conmutación y de cascada son las más adecuadas a la temática turística. Por un lado, la técnica de conmutación permite que, dependiendo del contexto en el que se encuentre el turista, se decida emplear un motor u otro. Por otro lado, cuando el número de puntos de interés es elevado, el uso de varios motores puede hacer más lenta la generación de las recomendaciones. En este caso, el uso de una técnica de hibridación en cascada es una buena opción para obtener resultados en menor tiempo.

II.2.1. Registro en el sistema

Para que un turista pueda recibir una recomendación, es necesario que se registre en el sistema y se identifique. Aunque el sistema posibilita el registro en una aplicación web, en este anexo se describen los procesos mediante la aplicación móvil.

Inicialmente, tendrá que indicar una serie de datos personales, como, por ejemplo, nombre de usuario, contraseña, datos personales, preferencias turísticas, etc.



Figura II.1. Acceso a RAMCAT

Al registrarse, la parte más relevante es la relacionada con el registro de las preferencias y gustos del turista. En la figura II.2, se puede observar cómo el turista puede indicar sus preferencias.

Se ha estimado más adecuado utilizar valores lingüísticos en vez de valores numéricos, ya las valoraciones subjetivas se ajustan mejor a este tipo de información.



Figura II.2. Detalle de registro de preferencias del usuario

II.2.2. Configuración de los motores de recomendación

Como se ha comentado anteriormente, el motor de recomendación es híbrido, pero dejamos al turista que, dependiendo de las circunstancias, pueda seleccionar una u otra forma de hibridación. Para que el turista no tenga que entrar en detalles se han definido varios modos de trabajo:

- Completa: Se le ha dado este nombre cuando todos los motores de recomendación trabajan simultáneamente. El turista podrá activar o desactivar los diferentes motores, estando por defecto todos activados.
- Optimizada: Esta opción utiliza el motor de recomendación basado en contenido sin memoria y demográfico. No tiene en cuenta visitas realizadas con anterioridad.

- Histórico: Esta opción sólo utiliza el motor de recomendación colaborativo y el basado en contenido a partir de las visitas realizadas.
- Cascada: Esta opción permite utilizar este modelo de hibridación, seleccionando el turista el orden de ejecución de los motores de recomendación. En la figura II.3 se puede observar la pantalla de ajustes, en la que además de indicar y configurar el modo de recomendación (su función esencial y fundamental), también puede configurarse el sistema para visitas grupales y atributos contextuales.



Figura II.3. Configuración de ajustes de recomendación

A continuación, se explican los diferentes modos de recomendación con mayor detalle.

II.2.2.1. Configuración de recomendación completa

Esta es la opción por defecto del sistema. La principal característica que presenta es que se ejecutan todos los motores de recomendación. En concreto, los modelos que utiliza son:

- a) Motor de recomendación basado en contenido sin memoria: como se indicó en el capítulo 6, este motor de recomendación se basa en las características del turista y en sus preferencias de planificación específicas, no teniendo en cuenta visitas anteriores realizadas.
- b) Motor de recomendación demográfico: en este motor no se tienen tampoco en cuenta las visitas realizadas previamente por el turista, obteniéndose una lista de puntos de interés que han sido bien valorados por turistas similares al que solicita la petición.
- c) Motor de recomendación colaborativo: en este motor se hace necesario que el turista haya realizado visitas previas, para que de esta forma puedan generarse recomendaciones precisas. En este módulo no se tienen en cuenta el contenido y las características de los ítems, sino que se buscan usuarios que hayan visitado los mismos que el turista que solicita la recomendación. Se recomendarán ítems que no haya visitado y que hayan sido valorados positivamente por dichos turistas.
- d) Motor de recomendación basado en contenido histórico: este motor es exactamente igual al motor del apartado a), pero en él se parte de las características de un ítem que ha sido bien valorado por el turista, buscando ítems que tengan características similares. Necesita por tanto que el turista haya realizado

previamente visitas y que seleccione el tipo de visita que desea realizar. En caso de no seleccionar ninguna, por defecto selecciona la de mayor calificación.

Una vez obtenidas las diferentes listas de preferencias, el modelo de hibridación de tipo mezcla las combina, pudiendo el turista activar o desactivar las recomendaciones de algún motor.

II.2.2.2. Configuración de recomendación optimizada

Esta opción tiene como principal beneficio que el número de cálculos a realizar para obtener la recomendación es menor, por lo que es una mejor opción desde el punto de vista computacional. Utiliza un método de hibridación de tipo cascada.

En primer lugar, se selecciona una lista de ítems con el motor basado en contenido sin memoria, siendo esta lista el conjunto de entrada del motor de recomendación demográfico. Su funcionamiento ha sido explicado en el capítulo 6.

Además, esta configuración es adecuada cuando se trate de un nuevo usuario, ya que los motores seleccionados no requieren información sobre visitas anteriores realizadas por el turista.

II.2.2.3. Configuración de recomendación histórica

En esta configuración se pretende realizar recomendaciones teniendo sólo en cuenta los datos históricos de visitas realizadas anteriormente por el turista. Utiliza dos motores de recomendación, uno es el motor de recomendación colaborativo y un segundo motor basado en contenido histórico. Se aplica la técnica de hibridación en cascada: en primer

lugar, el turista selecciona alguna de las visitas realizadas anteriormente que represente sus necesidades, obteniendo mediante el motor de recomendación basado en contenido histórico la lista de puntos de interés similar al seleccionado. Posteriormente, esta lista es el conjunto de entrada al motor de recomendación colaborativo, por lo que se obtiene una lista de puntos de interés ordenados a partir de las características de una visita satisfactoria anterior del turista y que haya sido bien valorada por otros.

II.2.2.4. Configuración de recomendación en cascada (Configuración avanzada)

Se trata de la opción más avanzada y requiere que el usuario haya utilizado habitualmente el sistema y entienda qué motores le pueden permitir obtener mejores recomendaciones.

En esta opción no sólo podrá indicar qué motor desea activar o no, sino también el orden de ejecución de los motores. En la figura II.4 se puede ver la manera en que el turista podrá configurar este modo de recomendación.



Figura II.4. Configuración avanzada de ajustes de recomendación

II.3. Actividades turísticas: información y evaluación

II.3.1. Información sobre las actividades turísticas

Como complemento al sistema de recomendación se debe ofrecer la posibilidad de que el turista pueda consultar información sobre el punto de interés que se le ha recomendado.

La posibilidad de mostrar diferentes tipos de información enriquece la experiencia del turista (información textual, fotos, videos, etc.). En las siguientes figuras se ilustra la forma en la cual el sistema ofrece información al turista sobre el recurso:

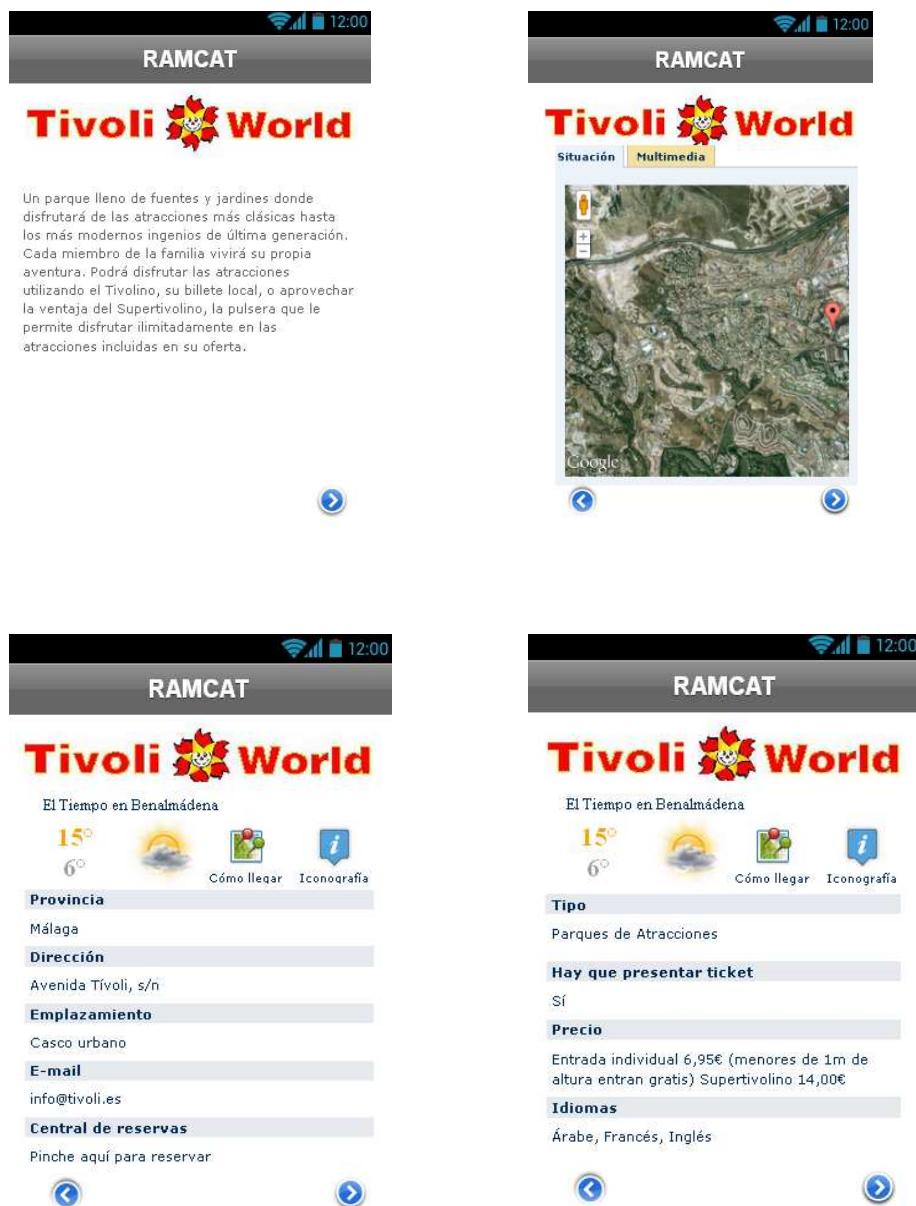


Figura II.5. Información sobre un punto recomendado

Como puede observarse en la figura II.5, el sistema ofrece una información completa sobre la actividad que le ha sido recomendada al turista.

II.3.2. Evaluación de las actividades turísticas

Una vez realizada una visita a un punto de interés, es necesario evaluarlo y que de esta forma el turista exprese su opinión sobre la actividad turística. En nuestro modelo, un punto de interés puede valorarse de dos formas: general y pormenorizada.

- a) Evaluación general: el turista debe evaluar en una escala lingüística lo que le ha parecido la visita realizada.
- b) Evaluación pormenorizada: el turista evalúa cada una de las características del ítem que ha visitado. Por ejemplo, si visitamos un restaurante, no sólo se evalúa la comida, también el lugar, tranquilidad, relación calidad/precio, servicio.... Aquellas características que no sean evaluadas por parte del turista se considerarán satisfactorias, no haciendo que sea un factor que destaque positiva ni negativamente sobre ese producto. Esta información será importante para el destino, pues modificará las valoraciones que los expertos han hecho sobre los puntos de interés. Otra posibilidad es dar una valoración por cada una de las diversas categorías a las que pertenezca el punto de interés.



Figura II.6. Valoración de una actividad recomendada

II.3.3. Características de las actividades turísticas.

Expertos turísticos

Cuando se desea añadir un nuevo punto de interés, el experto debe cumplimentar información sobre su tipo dentro de la taxonomía a la que pertenece, la valoración de diferentes características y el tipo de turista que visitará ese punto turístico.

Dado que el número de atributos puede ser muy diferente para cada tipo, existen características que no son aplicables a todos los puntos de interés. Obviamente, la característica “calidad comida” sólo tiene sentido en un restaurante y no en un museo.

Todo experto del destino turístico debe determinar el tipo de turista que suele visitar cada punto de interés. Para ello se utiliza una descripción lingüística, que puede ser consultada por el turista, pero no modificada, tal como aparece en las siguientes figuras:



Figura II.7. Valoración de un experto de una actividad

Además de esta información, se deben indicar las características del punto de interés y su tipo, así como un valor que indica lo relacionado que está con ese tipo. Por ejemplo, el recurso “Tivoli World” se considera como un recurso turístico relacionado con los tipos Ocio, parque de atracciones y parque de ocio. En el caso de describir características como parque de atracciones, el experto deberá indicar para cada atributo o propiedad su valor.

Por ejemplo, para parque de atracciones se consideran atributos o características como:

- a) Tipo de espacio.
- b) Ambiente.
- c) Edad objetivo.
- d) Precio.
- e) Lugares para descansar.
- f) Lugares para comer.

Otras características son contextuales:

- a) Acceso para minusválidos.
- b) Clima.
- c) Compañía.

En este caso el experto utiliza una escala entre 0 y 1, que posteriormente para poder ser comparada con valores lingüísticos del turista deben trasladarse a la escala correspondiente (Tabla II.1).



Figura II.8. Asignación de categoría a una actividad

En el siguiente ejemplo se muestran posibles valoraciones de cuatro características de varios puntos de interés relacionados con la categoría Ocio:

	Tipo de espacio (Característica F1)		Edad objetivo (Característica F2)				
	F1 _a	F1 _b	F2 _a	F2 _b	F2 _c	F2 _d	F2 _e
	Abierto	Cerrado	0-5 años	5-16 años	16-25	25-50	50-100
Tívoli	0,9	0,2	0,7	1	0,8	0,5	0,1
Sealife	0,2	0,9	0,4	0,8	0,7	0,7	0,7
Mariposario	0	1	0,4	0,7	0,3	0,6	0,9
Bioparc	0,9	0,2	0,7	0,7	0,3	0,5	0,5
Casino Marbella	0	0	0	0	0,2	0,9	1
Club Hielo	0	1	0	0,6	0,8	0,6	0,1
Aquavista	1	0	0,4	0,6	0,8	0,9	0,9
Kite School	0,9	0,2	0	0	0,9	0,6	0,1

	Ambiente (Característica F3)				Precio (Característica F4)		
	F3 _a	F3 _b	F3 _c	F3 _d	F4 _a	F4 _b	F4 _c
	Tranquilo	Animado	Familiar	Llamativo	Económico	Normal	Caro
Tivoli	0,3	0,8	0,9	0,8	0,2	0,7	0,3
Sealife	0,4	0,5	0,8	0,2	0,2	0,8	0,4
Mariposario	0,7	0,2	0,7	0,6	1	0,2	0
Bioparc	0,4	0,7	0,8	0,8	0,2	0,7	0,4
Casino Marbella	0,4	0,8	0	0	0	0,8	0,9
Club Hielo	0,2	0,9	0,7	0,5	0,6	0,8	0,3
Aquavista	0,8	0,3	0,8	0,7	0	0,4	0,9
Kite School	0	1	0,2	0,9	0,2	0,6	0,8

Tabla II.1. Características de distintas actividades

II.3.4. Características de las actividades turísticas. Turistas y expectativas

Como se ha descrito con anterioridad, algunos motores de recomendación utilizan técnicas basadas en contenido, las cuales necesitan que el turista especifique qué características deben tener los puntos de interés que desea visitar.

Dado que el número de atributos que puede tener un punto de interés es elevado, la introducción de información por parte del turista puede ser tediosa. No obstante, existen mecanismos para obtener algunas de estas características (no todas tienen que ser registradas explícitamente), reduciendo de esta manera la información que el turista indica sobre qué punto de interés desea visitar.

Una de los elementos de información necesarios es el tipo en el que se autoclasifica el turista (Figura II.9). En el registro, el turista informa sus características, pudiendo modificarlas más adelante:



Figura II.9. Descriptores turísticos del turista

Posteriormente, es posible realizar una búsqueda por categoría de punto de interés, pudiendo de esta forma refinar la búsqueda:

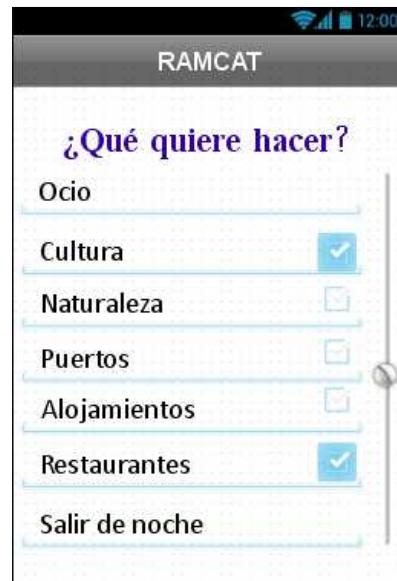


Figura II.10. Filtrado manual por parte del turista

Si selecciona varias categorías, el turista tendrá que indicar aquellas características que considera imprescindible en la recomendación. Si para una categoría existen diez características y sólo informa de dos, se interpretará que el resto de categorías le resulta indiferente. Si no desea seleccionar categorías, el turista puede seleccionar aquellas características que le interesen y el número que considere más adecuado. Naturalmente, cuanto mayor sea el número de características, más precisa será la recomendación.

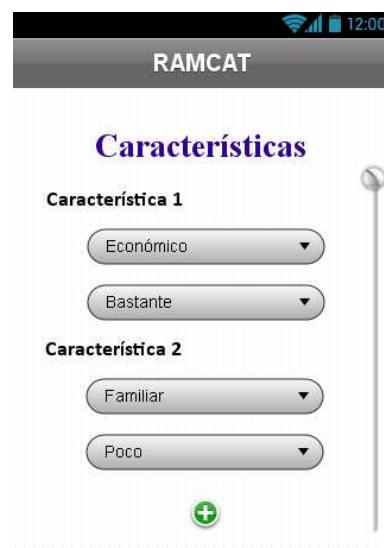


Figura II.11. Selección de características de productos por el turista

Otra forma de solicitar la recomendación, es seleccionando una o varias visitas ya realizadas (véase la figura II.12). De esta forma las características de los ítems recomendados serán las de los que ha seleccionado. En caso de que seleccione varios ítems, sólo se tendrán en cuenta para la recomendación las características comunes a todos ellos, asignándole a las no comunes el valor indiferente.



Figura II.12. Selección de productos visitados

II.4. Recomendaciones para grupos y generación de rutas

II.4.1. Creación de grupos

Una situación muy habitual es que varios turistas viajen juntos. La posibilidad de generar recomendaciones grupales hace necesaria la identificación y registro de grupos. Para crearlos, los diferentes usuarios del grupo deben estar registrados en el sistema. El usuario que cree un grupo tendrá el rol de administrador. Cuando éste solicite recomendaciones grupales, los resultados que se generarán tendrán en consideración las características de todos los miembros del grupo.

El administrador del grupo tendrá la posibilidad de añadir usuarios. Si un usuario que ha sido añadido al grupo no quiere pertenecer a ese grupo, tiene la opción de abandonarlo. Un usuario que no es administrador no puede añadir nuevos usuarios al grupo, sólo sugerirlos.

En la figura II.13, podemos observar la lista de turistas que forman un grupo:



Figura II.13. Lista de miembros de un grupo

En adelante, todas las recomendaciones se realizarán al grupo o turista virtual formado por todos los turistas del grupo. El modo de recomendación grupal requiere que todos los usuarios, una vez añadidos al grupo, den su consentimiento. En caso contrario no se generarán las recomendaciones grupales.

Un usuario puede pertenecer a varios grupos, pero cuando solicita recomendaciones grupales, ha de seleccionar el grupo, ya que no pueden proporcionársele recomendaciones que satisfagan simultáneamente a dos grupos.

II.4.2. Generación de rutas

Cuando el turista solicita una recomendación, el sistema de realidad aumentada mostrará en el dispositivo móvil los cinco puntos que se le recomiendan en orden de preferencias (Figura II.14).



Figura II.14. Imagen que visualiza recomendaciones

Para la generación de rutas el usuario dispone de dos opciones:

- a) Seleccionar manualmente los puntos que se desean visitar para generar una ruta (Figura II.15).



Figura II.15. Selección de puntos para generar ruta

- b) Indicar el tiempo que tiene para realizar la visita, dinero del que dispone, el primer punto de interés que desea visitar y si necesita tiempo para comer (Figura II.16). La generación de la ruta teniendo en cuenta estas características se realizará aplicando el modelo descrito en el capítulo seis.



Figura II.16. Parámetros de configuración de ruta

II.5. Conclusiones

En el presente anexo se ha presentado el prototipo de sistema de realidad aumentada móvil contextual aplicado al turismo. Como ya se ha indicado, se trata de un sistema de recomendación híbrido que permite, además, la generación de rutas y las recomendaciones para grupos.

Utiliza varios motores de recomendación basado en las siguientes técnicas:

- a) Colaborativo
- b) Demográficos
- c) Basados en contenido

La existencia de variantes y la posibilidad de realizar recomendaciones sin necesidad de visitas anteriores, además de tener en cuenta atributos contextuales hace del modelo propuesto una importante innovación en los destinos turísticos.

En este anexo hemos presentado las funcionalidades básicas del prototipo orientado al destino turístico de la Costa del Sol. Las principales aportaciones y característica del prototipo son:

- a) Ofrece un sistema de recomendación que permite al turista aplicar diferentes técnicas de hibridación, lo que permite hacer que la generación de recomendaciones sea más flexible que en los existentes.
- b) Ofrece a los turistas la posibilidad de utilizar atributos lingüísticos para describir sus necesidades y valoraciones,

facilitando la introducción de información cualitativa con expresiones cercanas y sencillas.

- c) Gestiona un modelo de adquisición de información de los turistas muy completo, pero a la vez poco exigente respecto a la cantidad de información necesaria para poder obtener recomendaciones.
- d) Permite generar recomendaciones sin necesidad de datos históricos sobre visitas anteriores. Al incorporar diferentes motores de recomendación, se permiten modos de configuración adaptados a diversas situaciones, como considerar visitas anteriores o no considerarlas porque no interesen o simplemente porque no existan.
- e) Permite la recomendación para grupos formados por varios turistas teniendo en cuenta las preferencias y necesidades de cada uno de ellos, lo que es una importante mejora en los sistemas de recomendación para destinos turísticos.
- f) Supone una importante alternativa para la mejora en la aplicación de nuevas tecnologías turísticas en el destino Costa del Sol.

Bibliografía

- Abowd, G.C., Atkeson, C.G., Hong, H., Long, S., Kooper, R. y Pinkerton, M. (1997): “Cyberguide: A mobile context-aware tourguide” Wireless Networks, 3(5):421-433
- Adomavicius,G. y Tuzhilin,A. (2005): “Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions.” IEEE Transaction on Knowledge and Data Engineering, 17(6):734–749, 2005
- Adomavicius, G., Sankaranarayanan, R., Sen, S. y Tuzhilin, A. (2005): “Incorporating contextual information in recommender system using a multidimensional approach”. ACM Transactions on Information Systems (TOIS), 23(1):103-145, 2005.
- Adomavicius,G., Tuzhilin,A., Berkovsky,S., y Said A.(2010): “Context-awareness in recommender systems: research workshop and movie recommendation challenge”. RecSys: 2010: 385 – 396
- Adomavicius, G., Mobasher, B., Ricci, F. y Tuzhilin, A. (2011): “Context-Aware Recommender Systems”. AI Magazine 32(3):67-80.
- Adomavicius, G. y Tuzhilin, A. (2011): “Context-aware recommender systems”. In Recommender Systems Handbook, pages 217-253.

Akrivas, G., Wallace, M., Andreou, G., Stamou, G. y Kollias, S. (2002):

“Context-sensitive semantic query expansions” In Proceedings of the IEEE International conference on artificial intelligence systems (IC AIS), pages 109-114, Divnomorskoe, Russia.

Arbib, M.A. (1995): “The handbook of brain theory and neural networks”. Cambridge.

Ardissono, L., Goy, G., Petrone, G., Segnan, M. y Torasso, P. (2003):

“Intrigue: personalized recommendation of tourist attractions for desktop and hand held devices” Applied Artificial Intelligence, 17(8):687-714.

Arfi, B. (2005): “Fuzzy decision making in politics: A linguistic fuzzy-set approach” Political analysis, 13(1):23-56

Armstrong, W.W. (1974): “Dependency structures of data base relationships” En IFIP Congress, pages 580-583.

Ashdown, D. (2012): “Mobile Search & Discovery”. Juniper Research.

Aukstakalnis, S., Blatner D. (1992): “Silicon Mirage. The Art and Science of Virtual Reality.” Peachpit Press, Berkeley, CA, USA, 1992.

Azuma, R. (1997): “A Survey of Augmented Reality”. Presence: Teleoperators and Virtual Environments 6, 4, pages 355-385.

Azuma, R.; Baillot, Y.; Behringer, R.; Feiner, S.; Julier, S. y MacIntyre, B. (2001): “Recent Advances in Augmented Reality”. IEEE Computer Graphics and Applications, Vol. 21, N6, pp 34-47

Balabanovic, M. y Shoham, Y. (1997): “Fab: Content-based collaborative recommender” Communications of the ACM, 40(3): 66-72

Baltrunas, L. y Ricci, F. (2009): “Context-dependent items generation in collaborative filtering” In Workshop on Context-Aware Recommender Systems (CARS 2009).

Barta, R., Feilmayr, C., Proll, B., Grun, C. y Werthner, H. (2009): “Covering the Semantic Space of Tourism, An approach based on modularized ontologies”. Proceedings of the 1st Workshop on Context, Information and Ontologies.

Bartfai, G. (1994): “Hierarchical Clustering with ART Neural Networks”. IEEE World Conference on Computational Intelligence, 940-944.

Bazire, M. y Brzillon, P. (2005): “Understanding context before using it” In A.Dey and et al., editors, Proceedings of the 5th International Conference on Modeling and Using Context. Springer-Verlag.

Belohlavek, R. (1999): “Fuzz Galois connections”. Mathematical Logic Quarterly, 45(4):497-504.

Belohlavek, R., Cordero, P., Enciso, M., Mora, A. y Vychodil, V. (2012): “An efficient reasoning method for dependencies over similarity and ordinal data”. In Torra, V., Narukawa, Y., López, B. y Villaret, M., editors, Modeling Decisions for Artificial Intelligence, volume 7647 of Lecture Notes in Computer Science, pages 408-419. Springer Berlin Heidelberg.

Berry, M.J. y Linoff, G. (1997): "Data mining techniques: for marketing, sales, and customer support" John Wiley & Sons, Inc. New York, USA

Bettman, J.R., Luce, M.F. y Payne, J.W. (1991): "Consumer decision making: A constructive perspective". In Tedeschi, M. (ed.) Consumer Behaviour and Decision Making, pp 1-42

Bezerra, B. y Carvalho, F. (2004): "A symbolic approach for content-based information filtering". Information Processing Letters, 92 (1)

Bhat, N. (1985): "Elements of applied stochastic processes". Wiley.

Billsus, D. y Pazzani, M.J. (2000): "User modeling for adaptive news access". User Modeling and User-Adapted Interaction, 10(2-3): 147-180

Bimber, O. y Raskar, R. (2005): "Spatial Augmented Reality. Merging Real and Virtual Worlds" Ed. A. K. Peters, Ltd.

Blanco, Y., Pazos, J.J., López, M., Gil, A. y Ramos, M. (2006): "AVATAR: An improved solution for personalized TV based on semantic inference". IEEE Transactions on Consumer Electronics, 52(1): 2236-232

Blanco, Y. (2007): "Propuesta metodológica para el razonamiento semántico en sistemas de recomendación personalizada y automática. Aplicación al caso de contenidos audiovisuales". Tesis doctoral. Universidad de Vigo.

Blum, A., Hellerstein, L. y Littlestone, N. (1995): "Learning in the presence of finitely or infinitely many irrelevant attributes". *Journal of Computer and System Sciences*, 50(1):32-40

Bobadilla, J., Ortega, F., Hernando, A. y Gutierrez, A. (2013): "Recommender systems survey". *Knowledge-Based Systems*, volumen 46, pages 109-132.

Bojadziev, G. y Bojadziev, M (1995): "Fuzzy sets, fuzzy logic, applications". World Scientific Publishing Ltd.

Brachman, R.J. y Schmolze, G. (1985): "An overview of the KL-ONE Know Representation system". *Cognitive Science*, 9(2):171-216

Breese, J.S., Heckerman, D. y Kadie, C. (1998): "Empirical analysis of predictive algorithms for collaborative filtering". In *Proceedings of the Fourteenth Conference on Uncertainty in Artificial Intelligence*, volume 461, pp 43-52.

Brown, P.J., Bovey, J.D. y Chen, X. "Context-aware applications: from the laboratory to the marketplace". *IEEE Personal Communications*, 4:58-64

Bueno, D. (2002): "Recomendación personalizada de documentos en sistemas de recuperación de la información basada en objetivos". Tesis doctoral. Universidad de Málaga.

Buhalis, D. y Law, R. (2008): "Progress in information technology and tourism management: 20 years on and 10 years after the Internet. The state of eTourism research". *Tourism Management*, 29

- Burke. R. (2000): “Knowledge-based recommender systems.” Encyclopedia of Library and Information Systems, 69(32)
- Burke, R, (2002): “Hybrid recommender systems: Survey and experiments”. User Modeling and User-Adapted Interaction, 12 (4):pp331–370
- Cantador, I. y Castells, P. (2009): “Semantic contextualisation in a news recommender system” In workshop on Context-Aware Recommender Systems (CARS 2009), New York, USA
- Cabrerizo, F.J. (2008): “Nuevos modelos de toma de decisión en grupo con información lingüística difusa”. Tesis Doctoral. Universidad de Granada.
- Caro, J.L. (2012): “Fotogrametría y modelado 3D: un caso práctico para la difusión del patrimonio y su promoción turística”. Turitec 2012.
- Castejón, R.y Méndez, E. (2012): “Introducción a la economía para turismo”. Editorial Prentice-Hall.
- Castillo, L.; Armengol, E.; Onaindía , E. y Sebastiá, L. (2008): “SAMAP: An user-oriented adaptative system for planning tourist visits”. Expert Systems with Applications, 34 : pp1318-1332
- Caudell, T.P., Mizell, D.W. (1992): “Augmented Reality: An Application of Heads-Up Display Technology to Manual Manufacturing Processes.” Proceedings of Hawaii International Conference on System Sciences, pp. 659-669, Hawaii, USA.

Cawood, S. y Fiala, M. (2008): “Realidad aumentada: A practical guide” Accesible desde <http://oreilly.com/catalog/9781934356036>

Celentano, M., Oliva, L., Vecchiio, P. y De Paolis, L (2011): “A simulation of life in a Medieval Town for edutainment and touristic promotion.” International Conference on Innovations in information Technology.

Cena, F., Console, L., Gena, C., Goy, A., Levi, G., Modeo, S. y Torre, I. (2006): “Integrating heterogeneous adaptation techniques to build a flexible and usable mobile tourist guide” AI Communications, 19(4):369-384.

Chao, D. L., Balthrop, J. y Forrest, S. (2005): Adaptive radio: achieving consensus using negative preferences”. En GROUP, pp 120-123.

Chen, D., Tsai, S., Hsu, C., Singh, J., Pal, G. y Girod, B. (2011): “Mobile augmented reality for books on a shelf”. In 2011 IEEE International Conference on Multimedia and Expo (ICME), pp.1-6, 2011.

Chen, Y.L., Cheng, L.C. y Chuang, C.N. (2008): “A group recommendation system with consideration of interactions among group members”. Expert Syst. Appl., vol. 34(3), pp 2082-2090

Cho, Y.H. y Kim, J.K. (2004): “Application of web usage mining and product taxonomy to collaborative recommendations in e-commerce” Expert Systems with Applications, 26(2):233-246.

Choudary, O., Charvillat, V., Grigoras, R. y Gurdjos, P. (2009): “MARCH: Mobile Augmented Reality for Cultural Heritage”. In

Proceedings of the 17th ACM international conference on Multimedia (MM '09). ACM, New York, NY, USA, 1023-1024

CIA (2012): “The CIA World Factbook”. Central Intelligence Agency, 2012.

Claypool, M., Gokhale, A., Miranda, T., Murnikov, P., Netes, D. y Sartin, M. (1999): “Combining content-bases and collaborative filters in an online newspaper”. In Proceedings of ACM SIGIR Workshop on Recommender Systems.

Colucci, S., Di Noia, T., Di Sciascio, E., Donini, F., Mongiello, M. y Motolla, M. (2003): “A formal approach to ontology-based semantic match of skills descriptions”. Journal of Universal Computer Science, 9(12):1437-1454.

Cordero, P., Enciso, M., Leiva, J.L., Mora, A. y Rossi, C. (2014): “Prefiltrado para sistemas de recomendación basado en lógica difusa”. XVII Congreso Español sobre Tecnologías y Lógica Fuzzy, Zaragoza, España.

De Carolis, B., Mazzota, I., Novielli, N. y Silvestri, V. (2009): “Using commom sense in providing personalized recommendations in the tourist domain”. In Workshop on Contex-Aware Recommender Systems (CARS 2009)

De Paolis, L.T., Aloisio, G., Celentano, M.G., Oliva, L. y Vecchio, P. (2011): “A simulation of life in a medieval town for edutainment

and touristic promotion". In 2011 International Conference on Innovations in Information Technology (IIT), pp. 361-366, 2011.

Deep Map (2001): "Deep Map: Intelligent, Mobile, Multi-Media and Full of Knowledge (Project Homepage)", European Media Laboratory,

<http://www.eml.org/english/research/deepmap/deepmap.html>

Deshpande, M. y Karypis, G. (2004): "Item-based top-n recommendation algorithms". ACM Transactions on Information Systems, 22(1):143-177.

Dey, A.K., Abowd, G.D. y Salber, D. (2001): "A conceptual framework and a toolkit for supporting the rapid prototyping of context-aware applications". Human-Computer Interaction, 16(2):97-166.

Dourish, P. (2004): "What we talk about when we talk about context". Personal and ubiquitous computing, 8(1):19-30.

du Boucher-Ryan, P. y Bridge, D. (2006): "Collaborative recommending using formal concept analysis". Knowledge-Based Systems, 19:309-315.

Dubois, D. y Prade,H. (1983): "Ranking fuzzy numbers in setting of possibility theory". Information Science 30: 183-224

Eirinaki, M. y Vazirgiannis, M. (2003): "Web mining for web personalization". ACM Transactions on Internet Technology, 3(1): pp 1-27

Esteve, R. y Fuentes, R. (2000): "Economía, historia e instituciones del turismo en España". Editorial Pirámide.

Euzénat, J., Remize, M. y Ochanine, H. (2003): "Projet Hi-Touch; Le web sémantique au secours du tourisme". Hélène Ochanine, W3C. Archimag.

Exceltur (2012): "Informe perspectivas turísticas". N°42. Octubre 2012

Fesenmaier, D.R., Ricci, F., Schaumlechner, E., Wober, K. y Zanella, C. (2003): "Dietorects: Travel advisory for multiple decision styles" In Information and Communication Technologies in Tourism 2003: Proceedings of the International Conference in Lausanne, 232-241.

Fesenmaier, D.R., Werthner, H., Wober, K. (2006): "Destination Recommendation System. Behavioural Foundations and Applications", CAB International, Oxfordshire-Cambridge.

Foni, A., Papagiannakis, G. y Magnenat-Thalmann, N. (2002): "Virtual Hagia Sophia: Restitution, Visualization and Virtual Life Simulation." In: UNESCO World Heritage Congress Proceedings, 2002.

Gago, A., Labandeira, X., Picos, F. y Rodríguez, M. (2009): "Specific and general taxation of tourism activities. Evidence from Spain". Tourism management, Volume 30, Issue 3, Pages 381-392.

Galindo, J., Urrutia, A. y Piattini, M. (2006): "Fuzzy Databases, Modeling, Design and Implementation". Ed. Idea Group Inc (IGI)

Ganter, B. y Wille, R. (1999): "Formal Concept Analysis". Springer-Verlag, 1999.

García, A., Torre, I. y Linaza, M.T. (2014): "Mobile social travel recommender system". Information and Communication Technologies in Tourism, pp 3-16

Gilbert, E. y Karahalios, K. (2009): "Predicting tie strength with social media" In Proceedings of the 27th international conference on Human factors in computing systems, pp 211-220. ACM, New York, USA.

Goh, D., Lee, C., Ang, R. y Lee, C. (2010); "Determining Services for the Mobile Tourist". The Journal of Computer Information Systems, 51 (1), 31-40.

Gruber, T.R. (1993): "A translation approach to portable ontologies". Knowledge Acquisition, 5(2):199-220.

Guevara, A., Aguayo, A. y Gálvez, S. (1997): "Internet y Turismo". Editorial Miramar.

Guevara, A. (2008): "Sistemas de información de destinos turísticos integrados (SIDTI)". I Jornada de investigación en turismo. Universidad de Sevilla.

Guevara, A., Aguayo, A, Gómez, I., Caro, J., Leiva, J. y otros (2009): "Sistemas informáticos aplicados al turismo". Editorial Pirámide.

Guo, G., Zhang, J. y Thalmann (2014): "Merging trust in collaborative filtering to alleviate data sparsity and cold start". *Knowledge-Based Systems*. Volume 57, pages 57-68

Guttman, R.H. (2006): "Merchant Differentiation through Integrative Negotiation in Agent-mediated Electronic Commerce". Tesis doctoral. School of Architecture and Planning, Massachusetts Institute of Technology.

Haarlsev, V. y Möller, R. (2001): "Racer system description". 1st International Joint Conference on Automated Reasoning, pp. 701-706.

Hájek, P. (2001). "On very true". *Fuzzy Sets and Systems*, 124(3):329-333.

Henderson, S. y Feiner, S. (2009): "Evaluating the benefits of augmented reality for task localization in maintenance of an armored personnel carrier turret". In IEEE International Symposium on Mixed and Augmented Reality (ISMAR), pp. 135-144.

Herlocker, J.L., Konstan, J.A., Terveen, L.G. y Riedl, J.T, (2004): "Evaluating collaborative filtering recommender systems". *ACM Transactions on Information Systems*, 22(1):5-53.

Herlocker, J.L. y Konstan, J.A. (2001): "Content.-independent task-focused recommendation" *IEEE Internet Computing*, pp 40-47, 2001.

Herrera, F., Herrera-Viedma y Martínez, L. (2000): “A fusion approach for managing multi-granularity linguistic term sets in decision making”. *Fuzzy Sets and Systems*, (114):43-58.

Herrero, G., Campo, A., Gil, M., García, A., Martin, D., Zugasti, I., Bilbao, S., Pérez, A. Koshutanski, H., Maña, A. y Pérez De Albeniz, I. (2012): “ConTur: An intelligent content management system for the tourism sector”. *Information and Communication Technologies in Tourism*, pp 36-47

Hinkle, D. y Toomey. C.N. (1994): “Clavier: Applying case-based reasoning on to composite part fabrication”. Proceeding of the 6th Application of AI Conference, Seattle, WA, AAAI Press, 55-62.

Holden, W. (2009): “Mobile Augmented Reality”. Juniper Research.

Holz, T., Campbell, A.G., O’Hare, G.M.P., Stafford, J.W., Martin,A. y Dragone, M. (2011): “MiRA. Mixed Realoty Agents”. *International Journal of Human-Computer Studies*, Volume 69, Issue 4, Pages 251-268.

Huang, C.L., Yeh, P.H., Lin, C.W. y Wu, D.C. (2014): “Utilizing user tag-based interests in recommender systems for social resource sharing websites”. *Knowledge-Based Systems*, volume 53, pages 86-96.

Hsu, C. (2011): “The Feasibility of Augmented Reality on Virtual Tourism Website”. In 4th International Conference on Ubi-Media Computing (U-Media), pp. 253-256.

IET. Instituto de estudios turísticos (2010): “Balance del Turismo. Resultados de la actividad turística en España. Año 2009”. Ed. Ministerio de Industria, Turismo y Comercio.

Isaacson, D. y Madsen, R. (1976): “Markov chains: Theory and Applications”. Wiley, 1976

Izkara, JL., Los Arcos, JL. y Maseda, JM. (2005): “Component-based approach for mixed reality applications development”. 7th Virtual Reality International Conference Laval, (VRIC). 20-24 Laval, France. 2005.

Jakkilinkki, R.; Georgievski, M. y Sharda, N. (2007): “Connecting Destinations with Ontology Bases e-Tourism Planner”. Information and Communication Technologies in Tourism. Ed. Springer, pp21-32.

Jameson, A. y Smyth, B.(2007): “Recommendation to groups”. The Adaptive Web, Methods and Strategies of Web Personalization , vol. 4321 Lecture Notes in Computer Science, pp 596-627. Springer.

Jensen, F.V. (2001): “Bayesian Networks and Decision Graphs”. Springer-Verlag.

Jin, R., Chai, J.Y. y Si, L. (2004): “An automatic weighting scheme for collaborative filtering”. Proceedings of the 27th annual international ACM SIGIR conference n Research and development in information retrieval, pp 3337-344.

- Jones, G.J.F., Glasnevin, D. y Gareth, I. (2005): "Challenges and opportunities of context-aware information access" In International Workshop on Ubiquitous Data Management, pp 53-62
- Juaneda, C.N. y Riera, A. (2011): "La oportunidad de la investigación en economía del turismo." Estudios de Economía Aplicada, 29(3), 711-722.
- Konstan, J.A., Riedl, J., Borchers, A. y Herlocker, J.L. (1998): "Recommender systems: A groupLens perspective". In AAAI Workshop Recommender Systems 98, Papers from the 1998 Workshop Technical report WS-98-08.
- Kramer, R.; Modsching, M. y Ten Hagen, K. (2007): "Development and evaluation of a context-driven mobile tourist guide". International Journal of Pervasive Computing and Communication(JPCC),3, 4, pp 378–399
- Kurbalija, V., Radovanovic, M., Geler, Z. y Ivanovic, M. (2014): "The influence of global constraints on similarity measures for time-series databases". Knowledge-Based Systems, volume 56, pages 49-67.
- Lamsfus, C., Alzua-Sorzabal, A., Martín, D., Salvador, Z., Usandizaga, A., (2009): "Contextual computing based Services". Tourism. Mediterranean Conference on Information Systems. Athens, Greece.
- Lamsfus, C., Grün, C., Alzua-Sorzabal, A. y Werthener, H. (2010): "Crear vínculos basados en el contexto para mejorar las experiencias de los turistas". Novática, nº203, 17-23

Lashkari, A., Parhizkar, B y Mohamedali, M.A. (2010): “Augmented Reality Tourist Catalogue Using Mobile Technology” Second International Conference on Computer Research and Development, 121-125

Lawrence, R.D., Almasi, G.S., Kotlyar, V., Viveros, M.S. y Duri, S.S. (2000): “Personalization of supermarket product recommendations.” IBM Research Division, T.J. Watson Research Center, Yorktown Heights, New York.

Le Grand, B., Soto, M. y Dodds, M. (2001): “XML Topic Maps and Semantic Web Mining”. XML Conference and Exhibition, pp 5-40

Lee, W.P., Kaoli, C., Huang, J.Y. (2014): “A smart TV system with body-gesture control, tag-based rating and context-aware recommendation”. Knowledge-Based System, Volume 56, pages 167-178.

Leiva, J.L., Caro, J.L., Guevara, A. y Arenas, M.A. (2008): “A cooperative method for system development and maintenance using workflow technologies” ICEIS (5) 2008: 130-135

Leiva, J.L., Enciso, M., Rossi, C., Cordero, P., Mora, A. y Guevara, A. (2013a): “Context-aware recommendation using Fuzzy Formal Concept Analysis”. In ICSOFT 2013, 8th International Joint Conference on Software Technologies.

Leiva, J.L., Guevara, A., Rossi, C. y Aguayo, A. (2012): “Sistemas de recomendación basados en grupo para su aplicación en realidad

aumentada”. Turitec 2012. Congreso Turismo y Tecnologías de la Información y las Comunicaciones (Málaga) pp 483-491

Leiva, J.L., Guevara, A. y Rossi, C. (2012): “Sistemas de recomendación para realidad aumentada en un sistema integral de gestión de destinos”. Revista de Análisis Turístico. N°14 pp 69-81

Leiva, J.L., Guevara, A., Rossi, C. y Aguayo, A. (2013b): “Sistemas de recomendación basados en grupo para su aplicación en realidad aumentada”. Revista Novática, n° 225, pp 63-67.

Leiva, J.L., Guevara, A., Rossi, C. y Aguayo, A. (2014): “Realidad aumentada y sistemas de recomendación grupales”. Estudios y perspectivas en Turismo, volumen 23, n°1, pp 40-59

Li, L. y Xu, Y. (2011): “Robust Depth Camera Based Eye Localization for Human-Machine Interactions”. Knowledge-Based and Intelligent Information and Engineering Systems, Lecture Notes in Computer Science, Springer Berlin / Heidelberg, 6881, pp. 424-435, 2011

Li, X. y Murata, T. (2010): “A Knowledge-based recommendation model utilizing formal concept analysis”. In 2nd International Conference on Computer and Automation Engineering (ICCAE 2010).

Liarokapis, F.(2006): “An exploration from virtual to augmented reality gaming”. Simulation & Gaming. SAGE Publications, Vol 37, 4, 507-533

Liburd, J. (2012): "Tourism research 2.0". *Annals of Tourism Research*, Volume 39, Issue 2, Pages 883-907

Lieberman, H., Dyke, N.W.V. y Vivacqua, A.S. (1999): "Let's browse: a collaborative browsing agent". *Knowledge Based System*, vol 12(8) pp 427-431

Linaza, M.T., Cobos, Y., Mentxaka, J., Campos, M.K. y Peñalba, M. (2007): "Interactive Augmented Experiences for Cultural Historical Events". In *Proceedings of VAST'2007*, pp.23-30.

Liu, D.R., Chen, Y.H. y Huang, C.K. (2014): "QAdocument recommendations for communities of question-answering websites". *Knowledge-Based Systems*, Volume 57, pages 146-160.

López, H., Navarro, A. y Relaño, J. (2010): "An analysis of augmented reality systems". Fifth International Multi-Conference on Computing in the Global Information, pp 245-250 (ICCGI 2010).

Lymberopoulos, D., Zhao, P., König, A., Berberich, K. y Liu, J. (2011): "Location-aware click prediction in mobile local search". CIKM, ACM, 2011, pp 413-422

Ma, J., Ruan, D., Xu, Y. y Zhang, G. (2007): "A fuzzy set approach to treat determinacy and consistency of linguistic terms in multi-criteria decision making". *International Journal of Approximate Reasoning*, 44(2):165-181.

Maamar, Z., Benslimane, D. y Narendra, N.C. (2006): “What can context do for web services?”. Communications of the ACM, 49(12):98-103

Mak, H., Koprinska, I. y Poon, J. (2003): “Intimate: a web-based movie recommender using text categorization”. In Proceedings of IEEE/WIC International Conference on Web Intelligence, 2003. WI 2003, pp 602-605

Maleszka, M., Mianowska, B. y Nguyen,N.T. (2013): “A method for collaborative recommendation using knowledge integration tools and hierarchical structure of user profiles”. Knowledge-Based Systems, Volume 47, pp 1-13.

Martinez, L, Pérez, L., Barranco, M y Espinilla, M. (2008); “Improving the effectiveness of knowledge based recommender systems using incomplete linguistic preference relations”. International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems.

McCarthy, J.F. “Pocket restaurant finder: A situated recommender systems for groups”. In Proceeding of Workshop on Mobile Ad-Hoc Communication at the 2002 Conference on Human Factors in Computer Systems.

McCarthy, K., Salamó, M., Coyle, L., McGinty, L., Smyth, B. y Nixon, P.(2006): “Cats: A synchronous approach to collaborative group recommendation”. En FLAIRS pp 86-91.

Mettouris, C. y Papadopoulos, G.A. (2014): “Ubiquitous recommender systems”. Computing, Volume 96, Issue 3, pp 223-257

Miele, A., Quintarelli, E., Rabosio, E. y Tanca, L. (2013): “A data-mining approach to preference-based data ranking founded on contextual information”. *Information Systems*, Volume 38, Issue 4, pp 524-544

Milgram, P. y Kishino, F. (1994): “A taxonomy of mixed reality visual displays” IEICE(Institute of Electronics, Information and Communication Engineers) Transactions on Information and Systems, Special issue on Networked Reality, Vol. 77, N. 12, pp. 1321-1329

Minguez, I., Berrueta, D. y Polo, L. (2009): “CRUZAR: an application of semantic matchmaking to eTourism”. *Cases on Semantic Interoperability Systems Integration: Practices and Applications*. Information Science Reference. Ed. IGI-Global.

Minsky, M. (1975): “A framework for representing knowledge”. *The Psychology of Computer Vision*, pp 211-277.

Mobasher, B. (2007): “Data Mining for web personalization”. *The adaptative web*, pp 90-135

Montaner, M., López, B. y De la Rosa, J. (2003): “A taxonomy of recommender agents on the Internet”. *Artificial Intelligence Review*, 19(4):285-330.

Mooney, R.J. y Roy, L. (2000): “Content-based book recommending using learning for text categorization” In Proceedings of the fifth ACM conference on Digital libraries, pages 195–204, 2000.

- Naciones Unidas (2008): “Cuenta satélite del turismo: Recomendaciones sobre el marco conceptual”. Departamento de Asuntos Económicos y Sociales. Estudios de métodos. Serie F, nº 80.
- Nasseri, H. (2008): “Fuzzy numbers: Positive and Nonnegative”. International Mathematical Forum, 3, 2008, nº36, 1777-1780
- Nenova, E., Ignatov, D.I. y Konstantinov, A.V. (2013): “An FCA-based Boolean matrix factorization for collaborative filtering”. In Proceedings of the Workshop Formal Concept Analysis Meets Information Retrieval (FCAIR 2013)
- Niu, L., Yan, X., Zhang, C. y Zhang,S. (2002):“Product hierarchy-based customer profiles for electronic commerce recommendation”. In Proceedings of 2002 International Conference on Machine Learning and Cybernetics, volume 2, pages 1075–1080, 2002
- Noguera, J.M., Barranco, M.J., Segura, R.J. y Martínez, L. (2012): “A Location-Aware Tourism recommender system based on mobile devices”. 10th International FLINS Conference on Uncertainty in Knowledge Engineering and Decision Making, pp 34-39
- Noy, N.F. y McGuinness, D.L. (2001): “Ontology Development 101: A guide to creating your first ontology”. Stanford Knowledge Laboratory Technical Report KSL-01-05.
- Oku, K., Nakajima, S., Miyazaki, J. y Uemura, S. (2006): “Context-aware SVM for content-dependent information recommendation”. In Proceedings of the 7th International Conference on Mobile Data Management, pp109-110
- O'Connor, M., Cosley, D., Konstan, J.A. y Riedl, J. (2001): “Polylens: a recommender systems for groups of users”. In ECSCW'01

- Proceedings of the 7th Conference on European Conference on Computer Supported Cooperative Work, pp199-218. Kluwer Academic Publishers, Norwell, MA, USA.
- Olwal, A., Lindfors, C., Gustafsson, J., Kjellberg, T. y Mattsson, L. (2005a): “ASTOR: An Autostereoscopic Optical See-through Augmented Reality System”. Proceedings Fourth IEE and ACM Internation Symposium on Mixed and Augmented Reality, pp 24-27
- OMT (2001): “Tesauro de Turismo y Ocio”. Organización Mundial del Turismo, 2001.
- OMT (2012): “Panorama OMT del turismo internacional. Edición 2012”. Organización Mundial del Turismo, 2012.
- Ou, S., Pekar, V., Orasan, C., Spurk, C. y Negri, M. (2008): “Development and alignment of a domain-specific ontology for question answering”. Proceedings of the 6th Edition of the Language Resources and Evaluation Conference.
- Olwal, A. y Höllerer, T. (2005b): “POLAR: Portable, Optical see-through, Low-cost Augmented Reality” In Proceedings of the ACM symposium on Virtual reality software and technology (*VRST '05*). ACM, New York, NY, USA, 227-230.
- Palmisano, C., Tuzhilin, A. y Gorgoglione, M. (2008): “Using context to improve predictive modeling of customers in personalization applications”. IEEE Transactions on Knowledge and Data Engineering, 20(11):1535-1549.

- Panniello, U., Tuzhilin, A., Gorgoglione, M., Palmisano, C. y Pedone, A. (2009): “Experimental comparison of pre-vs, post-filtering approaches in context-aware recommender systems”. In Proceedings of the 3rd ACM Conference on Recommender systems, pp265-268.
- Pasman, W., Woodward, C., Hakkarainen, M., Honkamaa, P., y Hyvakkä, J. (2004): “Augmented Reality with Large 3DModels on a PDA – Implementation, Performance and Use Experiences”. Int Conf on Virtual-Reality Continuum & Its Applications in Industry (VRCAI,Singapore, June 16-18), 344-351
- Pazzani, M.J. (1999): “A framework for collaborative, content-based and demographic filtering”. Artificial Intelligence Review, 13(5-6):393-408, 1999
- Pérez, D. (2009): “Desarrollo de sistemas de realidad virtual y aumentada para la visualización de entornos acrofóbicos. Estudios comparativos entre ellos.” Tesis doctoral. Universidad de Valencia.
- Pérez, L.G. (2008): “Modelo de recomendación con falta de información. Aplicaciones al sector turístico”. Tesis doctoral. Universidad de Jaén.
- Piliponyte, A. (2012): “Sequential Group Recommendations”. MA Thesis, Free University of Bozen–Bolzano, 2012.
- Porcel, C., Herrera-Viedma, E. (2010): “Dealing with incomplete information in a fuzzylinguistic recommender system to disseminate information in university digital library”. Knowledge-Based System, 23, 32-39

Portalés, C., Lerma, J. L., y Pérez, C. (2009): “Photogrammetry and augmented reality for cultural heritage applications”. *The Photogrammetric Record*, 24 (128), 316-331.

Prahalad, C.K. (2004): “Beyond CRM: CK Prahalad predicts customer context is the next big thing”. American Management Association MwWorld.

Resnick, P. y Varian, H.R. (1997): “Recommender systems. Association for Computing Machinery”. *Communications of the ACM.*, 40(3):56, Mar 1997.

Ruiz, A., Urdiales, C., Fernández-Ruiz, J.A. y Sandoval, F. (2004): “Ideación arquitectónica asistida mediante realidad aumentada”. *Actas de XIV Jornadas Telecom I+D*, pp. 101-109.

Ryan, N., Pascoe, J. y Morse, D. (1997): “Enhanced Reality Fieldwork: the context-aware archaeological assistant”. Gaffney, V., van Leusen, M. y Exxon, S. (eds) *Computer Applications in Archaeology*. British Archaeological Reports, Oxford.

Rodrigues, J.A., Cardoso,L.F., Moneria, J. y Xexéo, G. (2013): “Bringing knowledge into recommender systems”. *Journal of Systems and Software*, Volume 86, Issue 7, pp 1751-1758.

Rodríguez, B., Molina, J. y Caballero, R. (2010): “Sistema de ayuda al turista para el diseño de un viaje personalizado”. *Actas Turitec 2010*, Málaga

Rodríguez, B., Molina, J., Pérez, F. y Caballero, R. (2012): "Interactive design of personalised tourism routes". *Tourism Management* 33 (2012):926-940

Rodríguez, R.M, Martínez, L. y Herrera, F. (2012): "Hesitant fuzzy linguistic term sets for decision making." *IEEE Transactions on Fuzzy Systems*. Volume 20, Issue 1, pages 109-119.

Sánchez, E. (1979): "Inverses of fuzzy relations, applications to possibility distributions and medical diagnosis". *Fuzzy Sets and Systems*, 2, 75-86.

Sarwar, B.M., Karypis, G., Konstan, J.A. y Reidl, J. (2001): "Item-based collaborative filtering recommendation algorithms". In Proceedings of the 10th International World Wide Web Conference (WWW10), pp 285-295.

Schafer, J., Frankowski, D., Herlocker, J. y Sen, S. (2007): "Collaborative Filtering Recommender Systems" Vol 4321, Chapter 9, pp 291-325. Heidelberg, S.B.

Schiaffino, S. y Amandi, A. (2009): "Building an expert travel agent as a software agent". *Expert Systems with Applications*, 36(2, Part 1):1291-1299.

Schilit, B.N. y Theimer, M.M. (1994): "Disseminating active map information to mobile hosts". *IEEE network*, 8(5):22-32.

Schiller, J.H. y Voisard, A. (2004): "Location-based services". Morgan Kaufmann, 2004.

Schwab I., Kobsa A. y Koychev I. (2001): "Learning user interests through positive examples using content analysis and collaborative

- filtering". Technical report, Fraunhofer Institute for Applied Information Technology.
- Sieg, A., Mobasher, B. y Burke, R. (2007): "Representing context in web search with ontological user profiles". In Proceedings of the 6th International Conference on Modeling and Using Context.
- Sigala, M. (2011): "Special Issue on Web 2.0 in travel and tourism: Empowering and changing the role of travelers". Computers in Human Behaviour, Volume 27, Issue 2, Pages 607-608.
- Siorpaes, K. y Bachlechner, D. (2006): "OnTour: Tourism Information Retrieval based on YARS". Proceedings of the European Semantic Web Conference.
- Siricharoen, W.V. (2010): "Enhancing semantic web and ontologies for e-tourism". International Journal of Intelligent Information and Database Systems, Volume 4, Issue 4, pages 355-372.
- Slater, M. y Wilbur, S. (1997): "A framework for immersive virtual environment (FIVE): Speculations on the role of presence in virtual environments." Presence: Teleoperators and Virtual Environments, Vol. 6, N. 6, pp. 603–616, 1997
- Seungjun, K. y Anind, K.V. (2010): "AR interfacing with prototype 3D applications based on user-centered interactivity". Computer-Aided Design, Volume 42, Issue 5, pp 373-386
- Smyth, B. y Cotter, P. (2000): "A personalized tv listings service for the digital tv age". Journal of Knowledge-Based Systems, 13(2-3):53-59.

- Smyth B. (2007): “Case-based recommendation” En The Adaptative Web, pages 342-376. Springer Berlin / Heidelberg.
- Staab, S. y Studer, R. (2004): “Handbook on Ontologies”. International Handbooks on Information Systems, Springer, 2004
- Steganidis, K., Pitoura, E. y Vassiliadis, P. (2007): “A context-aware preference database system”. International Journal Pervasive Computing and Communications, 3(4):439-600
- Stricker, D., Pagani, A. y Zoellner, M. (2009): “In-Situ Visualization for Cultural Heritage Sites using Novel Augmented Reality Technologies”. In International meeting on graphic archeology and informatics, cultural heritage and innovation, pp. 152-158.
- Suárez-Figueroa, M.C. y Gómez-Pérez, A. (2008): “First attempt towards a standard glossary of ontology engineering terminology”. Proceedings 8th International Conference on Terminology and Knowledge Engineering.
- Suárez-Figueroa, M.C. y Gómez-Pérez, A. (2009): “NeOn: Methodology for building ontology networks: a scenario-based methodology”. Proceedings of the International Conference on Software, Services and Semantic Technologies, pages 160-167
- Sundbo, J., Orfila, S. y Sorensen, F. (2007): “The innovative behaviour of tourism firms. Comparative studies of Denmark and Spain”. Research Policy, n°88, pp. 88–106

- Tejada-Lorente, A., Porcel, C., Peis, E., Sanz, R. y Herrera-Viedma, E. (2014): “A quality based recommender system to disseminate information in a university digital library”. *Information Sciences*, volume 261, pages 52-69.
- Terveen, L. y Hill, W. (2002): “Human-Computer Interaction in the New Millennium” *Human-Computer Collaboration in Recommender Systems*. Addison-Wesley, 2002.
- Thomas, B., Close, B., Donoghue, J., Squires, J., De Bondi, P., Morris, M. y Piekarski, W. (2000): “ARQuake: An Outdoor/Indoor Augmented Reality First Person Application.” *Proceedings of the 4th International Symposium on Wearable Computers*, pp. 139-146, 2000
- Thomas, K. y Kilmann, R. (1974): “Thomas-Kilmann Conflict Mode Instrument”. Tuxedo, New York, 1974
- Troitiño, L, Rodriguez, A. y Hidalgo, C. (2011): “Nuevas tecnologías aplicadas al turismo. Una aproximación a la realidad aumentada”. *Revista de investigación de la Universidad a distancia de Madrid*, Volumen 1, Año 2011
- UNWTO (2011): “Tourism Towards 2030. Global Overview”. World Tourism Organization (UNWTO), 2011.
- Uschold, M. y Grüninger, M. (1996): “Ontologies: principles, methods and applications”. *Knowledge Engineering Review*, 11(2),93-155
- Valdés, L., Valle, E. y Sustacha, I. (2011): “El conocimiento del turismo en el ámbito regional.” *Cuadernos de Turismo*, 27, pp 931–952

Valentini, P. (2009): “Interactive virtual assembling in augmented reality”. International Journal on Interactive Design and Manufacturing. Vol. 3, pp. 109–119.

Valls, J. (2003). “Las Claves del Mercado Turístico. Cómo Competir en el Nuevo Entorno” Bilbao: Deusto.

Van Krevelen, D.W.F. y Poelman, R. (2010): “A Survey of Augmented Reality Technologies, Applications and Limitations.” The International Journal of Virtual Reality, Vol. 9, No. 2, pp. 1-20.

Van Setten, M., Pokraev, S. y Koolwaaij, J. (2004): “Context-aware recommendations in the mobile tourist application compass”. In W.Nejdl and P. De Bra, editors, Adaptive Hypermedia, pp 235-244. Springer Verlag, 2004.

Visser, A. (2011): “Survey of XML Languages for Augmented Reality Content” In Proceedings of AR Standardization Forum, Barcelona, pp. 52-59, 2011.

Vlahakis, V., Ioannidis, M., Karigiannis, J. , Tsotros, M., Gounaris, M. , Stricker, D., Gleue, T., Daehne, P. y Almeida, L. (2002): “Archeoguide: an augmented reality guide for archaeological sites”. IEEE Computer Graphics and Applications. Vol. 22, No. 5, pp. 52-60, 2002.

Vlahakis, V., Demiris, T. y Ioannidis, N. (2004) : “LIFEPLUS Cultural heritage dissemination on a wide range of client devices: from the simple handheld to the advanced AR platform.” Multi-Platform e-Publishing, 2004.

Vozalis, M.G. y Margaritis, K.G. (2007): "Using SVD and demographic data for the enhancement of generalized collaborative filtering". *Information Sciences*, 177(15):3017-3037, 2007.

Wagner, D. y Schmalstieg, D. (2003): "First steps towards handheld augmented reality." *Proceedings of the 7th IEEE International Symposium on Wearable Computers (ISWC 03)*, pp. 127-135, 2003.

Walderhaug, S., Stav, E., Johansen, U. y Olsen, G. (2009): "Traceability in Model-Driven Software Development". *Designing Software-Intensive Systems: Methods and Principles*. IGI Global, pp 133-159.

Wang, J.Y. y Kao, H.Y. (2013): "RSOL: A trust-based recommender system with an opinion leadership measurement for cold start users". *Information Retrieval Technology*, volume 8201, pp 500-512.

Wang, Y.M., Yang, J.B. y Xu, D.L. (2005): "A preference aggregation method through the estimation of utility intervals". *Computers and Operations Research*, 32, pp 2027-2049

WebGuide (2001): "WebGuide: A City Guide for the Internet". European Media Lab,
<http://www.eml.org/english/research/deepmap/deepgis/webguide.html>

Weigand, L. (1997): "A multilingual ontology-based lexicon for news filetering – The TREVI project". *15th International Joint Conferences on Artificial Intelligence*, pp 160-165.

Wille, R. (1982): "Restructuring lattice theory: an approach based on hierarchies of concepts". Ordered Sets, pages 445-470.

Wille, R. (1989): "Knowledge Acquisition by Methods of Formal Concept Analysis", In E. Diday, editor, Data Analysis, Learning Symbolic and Numeric Knowledge, pp. 365-390

Winston, W.L. (2005): "Investigación de operaciones". Cengage Learning Latin America.

Wolf, J., Aggarwal, C., Wu, K.L. y Yu, P. (1999): "Horting hatches an egg: A new graph-theoretic approach to collaborative filtering". Proceedings of ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Diego, CA., pp 201-212

Xiang, Z. y Pan, B. (2011): "Travel queries on cities in the United States: Implications for search engine marketing for tourist destinations." Tourism Management, 32, p.88-97.

Xu, Z.S.(2004): "A method based on linguistic aggregation operators for groupdecision making with linguistic preference relations". Information Science,166:19–30, 2004

Yang, R. (2011): "The study and improvement of Augmented reality based on feature matching". In IEEE 2nd International Conference on Software Engineering and Service Science (ICSESS), pp. 586-589.

Yeo, C.T., Ungi, T., U-Thainual, P., Lasso, A., McGraw, R.C. y Fichtinger, G. (2011): "The Effect of Augmented Reality Training on Percutaneous Needle Placement in Spinal Facet Joint

- Injections". In IEEE Transactions on Biomedical Engineering. Vol. 58, No. 7, pp. 2031-2037.
- Younghoom, K. y Kyuseok, S. (2014): "TWILITE: A recommendation system for Twitter using a probabilistic model based on latent Dirichlet allocation". Information Systems, Volume 42, pp 59-77.
- Zadeh, L.A. (1996): "Fuzzy logic = computing with words". IEEE Transactions on Fuzzy Systems, 4(2):103-111.
- Zenebe, A. y Norcio, A.F. (2009): "Representation, similarity measures and aggregation methods using fuzzy sets for content-based recommender systems". Fuzzy Sets and Systems, 160:76-94.
- Zeng, C., Xing, C.X., Zhou, L.Z. y Zheng, X.H. (2004): "Similarity measure and instance selection for collaborative filtering". International Journal of Electronic Commerec, 8(4):115-129
- Zhang, Y.,Chen, W. y Yin, Z. (2013): "Collaborative filtering with social regularization for TV program recommendation". Knowledge-Based Systems, volume 54, pages 310-317.
- Zhang, Z., Cui, P. y Cui, H. (2006) : "Recovery of Egomotion from Optical Flow with Large Motion Based on Subspace Method". In IEEE International Conference on Robotics and Biomimetics, pp. 555-560, 2006.
- Zoellner, M., Keil, J., Drevensek, T. y Wuest, H. (2009): "Cultural Heritage Layers: Integrating Historic Media in Augmented Reality". 15th International Conference on Virtual Systems and Multimedia, pp. 193-196.

Zou, Y., Finin, T., Ding, L., Chen, H. y Pan, R. (2003): "Using Semantic Web Technology in Multi-Agent Systems: a Case Study in the TAGA Trading Agent Environment". Proceedings of the 5th international conference on electronic commerce, pp 95-101.

Índice de figuras

2.1. Relación entre realidad virtual y realidad aumentada.....	28
2.2. Video see-through y optical see-through	31
2.3. Visor panorámico	38
2.4. Funcionamiento del sistema POLAR	44
2.5. Funcionamiento del sistema Layar.....	45
2.6. Esquema de elementos comunes a los navegadores de RA	48
2.7. Audio guías. Cortometrajes y TouchTech Tourist Guide.....	52
2.8. Proyecto Lifeplus	56
2.9. Representación virtual del Templo de Hera.....	56
2.10. Representación del Palacio de Diana (Italia)	58
2.11. Dispositivos para aplicaciones orientadas al guiado	58
2.12. Uso de Wikitude en un smartphone	60
2.13. Uso de Layar en un smartphone	61
2.14. Uso de Junao en funcionamiento.....	63

2.15. Uso de Mixare en funcionamiento	64
2.16. Uso de contenido textual en realidad aumentada.....	66
2.17. Objeto 3D de realidad aumentada en publicidad	67
3.1. Estructura jerárquica de información contextual	121
3.2. Modelo multidimensional	122
3.3. Componentes en un proceso de recomendación general	127
3.4. Paradigmas contextuales en SR	128
5.1. Arquitectura propuesta modelo RAMCAT	176
5.2. Componentes de la arquitectura RAMCAT	177
5.3. Pantalla del prototipo móvil RAMCAT	185
5.4. Visión parcial de la ontología desarrollada para RAMCAT	206
5.5. Tipos de conexiones en la taxonomía propuesta.....	210
5.6. Visión parcial de la taxonomía correspondiente al punto de interés Catedral de Málaga	210
5.7. Detalle aplicación indicando características contextuales.....	239
5.8. Preferencias específicas del usuario u	246
5.9. Motores de recomendación de la propuesta	250
5.10. Evolución del descriptor turístico d1 en el usuario u1	262
5.11. Evolución del descriptor turístico d2 en el usuario u1	262
5.12. Evolución del descriptor turístico d3 en el usuario u1	263

6.1. Reducción del número de ítems mediante el proceso PFCBI	268
6.2. SR contextual basado en implicaciones difusas	272
6.3 Herramienta diseñada para simular el motor PFCBI.....	289
6.4. Comparativa ítems que son pre-filtrados y rechazados por el motor PFCBI	289
6.5. Función de pertenencia triangular	299
6.6. Función de pertenencia trapezoidal.....	299
6.7 Función de pertenencia gaussiana.....	300
6.8. Función de pertenencia pseudo-exponencial.....	300
6.9. Escalas lingüísticas en un contexto multigranular.....	302
6.10. Ejemplo de varias funciones de pertenencia	302
6.11 Medida de posibilidad	303
6.12. Medida de necesidad	304
6.13. Escalas lingüísticas utilizadas en el ejemplo	313
6.14. Significado semántico de las escalas utilizadas en el ejemplo	314
6.15. Escalas lingüísticas y significado semántico.....	338
6.16. Destino con diez ítems y diez calificaciones por ítem	346
6.17. Destino con diez ítems y cien calificaciones por ítem	346
6.18. Destino con cien ítems y cien calificaciones por ítem	347

6.19. Destino con cien ítems y mil calificaciones por ítem.....	347
6.20. Destino con cien ítems y 10000calificaciones por ítem	347
6.21. Destino con diez ítems y diez calificaciones por ítem	348
6.22. Destino con diez ítems y cien calificaciones por ítem	348
6.23. Destino con cien ítems y cien calificaciones por ítem	349
6.24. Destino con cien ítems y mil calificaciones por ítem.....	349
6.25. Destino con cien ítems y 10000 calificaciones por ítem.....	349
6.26. Destino con mil ítems y 10000 calificaciones por ítem.....	350
6.27. Escala lingüística y significado semántico aplicado en el ejemplo..	364
6.28. Esquema recomendación grupal	391
6.29. Escalas lingüísticas utilizadas en el ejemplo	403
6.30. Significado semático de las escalas utilizadas en el ejemplo.....	404
6.31. Grafo dirigido sobre tres puntos turísticos	414
6.32. Predicción de interés de los ítems en cada motor de recomendación	
.....	433
II.1. Acceso a RAMCAT.....	463
II.2. Detalle de registro de preferencias del usuario	464
II.3. Configuración de ajustes de recomendación.....	465
II.4. Configuración avanzada de ajustes de recomendación.....	469

II.5. Información sobre un punto recomendado.....	470
II.6. Valoración de una actividad recomendada	471
II.7. Valoración de un experto de una actividad	472
II.8. Asignación de categoría a una actividad	474
II.9. Descriptores turísticos del turista.....	476
II.10. Filtrado manual por parte del turista.....	477
II.11. Selección de características de productos por el turista	477
II.12. Selección de productos visitados.....	478
II.13. Lista de miembros de un grupo	479
II.14. Imagen que visualiza recomendaciones	480
II.15. Selección de puntos para generar ruta.....	481
II.16. Parámetros de configuración de ruta.....	481

Índice de tablas

2.1. Diferencias entre realidad virtual y realidad aumentada.....	27
2.2. Componentes hardware de un sistema de realidad aumentada	35
2.3. Requisitos software en un sistema de realidad aumentada.....	36
2.4. Características técnicas de ULTRA	41
2.5. Características técnicas de MARCH	42
2.6. Características técnicas de ASTOR	43
2.7. Características técnicas de POLAR	44
2.8. Características técnicas de LAYAR	46
2.9. Características técnicas de LOCUS.....	47
2.10. Características de contenidos en aplicaciones de RA.....	68
3.1. Técnica de recomendación colaborativa	86
3.2. Técnica de recomendación basada en contenido.....	87
3.3. Técnica de recomendación demográfica.....	88

3.4. Técnica de recomendación basada en conocimiento	89
3.5. Valoraciones de turistas	99
3.6. Valoraciones de un turista sobre museos.....	103
5.1. Descripción de características de tipo de espacio	193
5.2. Clasificación de categorías turísticas.	195
5.3. Clasificación de subcategorías de la categoría Monumentos y Patrimonio.....	196
5.4. Ejemplo de valoraciones en la Catedral de Málaga	213
5.5. Valoración de atributos no contextuales	217
5.6. Ejemplo de valoraciones sobre preferencias de dos turistas.....	223
5.7. Necesidades de un usuario que solicita restaurante	227
5.8. Ejemplo de restricciones de un usuario	247
5.9. Valor de descriptores turísticos de un grupo de turistas	259
5.10. Valor de descriptores turísticos de actividades recomendadas....	260
5.11. Valoración de cada turista en cada actividad visitada	260
5.12. Comparación retroalimentación del perfil de los descriptores turísticos	261
6.1. Representación en AFC de unos restaurantes y su diseño	271
6.2. Lista de ítems.....	282
6.3. Representación en AFC de los ítems y sus características	282

6.4. Descripción de puntos turísticos.....	314
6.5. Descripción de necesidades del usuario	315
6.6. Grado de desempate de los ítems	318
6.7. Ejemplo de estereotipos	328
6.8. Descripción de puntos turísticos.....	340
6.9. Descripción contextual y estereotipo del usuario.....	340
6.10. Descriptores turísticos-demográficos.....	342
6.11. Valoración de predicción de interés de los ítems	344
6.12. Actividades visitadas por el turista que solicita recomendación .	362
6.13. Actividades visitadas por los turistas amigos	363
6.14. Actividades vecinas	364
6.15. Descripción contextual de los puntos turísticos	365
6.16. Descripción contextual del usuario.....	365
6.17. Tabla de similitudes entre ítems y valor medio de ítems	367
6.18. Tabla de medias ajustadas	368
6.19. Tabla de sumas ponderadas	369
6.20. Tabla de predicción	370
6.21. Visitas realizadas por turista que solicita recomendación.....	380
6.22. Tabla de pesos normalizados para cada visita.....	383

6.23. Tabla de valores k_c^i	384
6.24. Lista de puntos turísticos para recomendar.....	404
6.25. Lista de usuarios que forman un grupo	405
6.26. Trazabilidad de visitas	413
6.27. Resumen de ítems recomendados	433
II.1. Características de distintas actividades	475

Abreviaturas

API.....	Application Programming Interface
AFC.....	Análisis formal de conceptos
ARML	Augmented Reality Markup Language
CARS	Sistemas de recomendación contextuales
CAVE.....	Sistema de habitación inmersiva
CRS.....	Sistema Central de Reservas
DMD	Digital Micromirror Device
ECC.....	European E-commerce Conference
FBCC	Filtrado basado en características contextuales
FBCCG	Filtrado basado en características contextuales grupales
FBSMG	Filtrado basado en contenido sin memoria grupal
FCA.....	Formal Concept Analysis
GDS.....	Sistema Global de distribución
GND.....	Grado de no dominancia
GPS.....	Global Positioning System (Sistemas de posición global)
GSM	Global System for Mobile communications
HMD.....	Head Mounted Display (Dispositivos acoplados a la cabeza)

HOE.....	Elementos holográficos ópticos transparentes.
KML	Keyhole Markup Language
LBE	Location Based Services
OMT	Organización mundial del turismo
ORSD.....	Documento de especificación de requerimientos de ontología
OWL	Web Ontology Language
PDA	Personal Digital Assistant
PFCBI.....	Pre-filtrado contextual basado en implicaciones
PIB.....	Producto Interior Bruto
PIT.....	Punto de interés turístico
POI	Point of interest
QR	Quick Response code
RA.....	Realidad aumentada
RDF	Resource Description Framework
RM	Realidad mixta
RSS	Really Simple Syndication (Sistema de redistribución)
SIG.....	Sistema de información geográfico
SCP ^u	Preferencias de especificación contextual del usuario <i>u</i>
SGP ^u	Preferencias específicas del grupo al que pertenece el usuario <i>u</i>
SHP ^u	Preferencias de especificación históricas del usuario <i>u</i>
SP ^u	Preferencias específicas del usuario <i>u</i>
SPP ^u	Preferencias específicas de planificación del usuario <i>u</i>
SRP ^u	Preferencias de recomendación específicas del usuario <i>u</i>
TAO	Topics, Associations & Ocurrences

TIC.....	Tecnologías de la Información y la Comunicación
UMPC	PC ultra móvil
VA.....	Virtualidad aumentada
WLAN	Wireless Local Area Network