



Método para sistemas de recomendaciones sensibles al contexto en entornos turísticos

Method of Context Aware Recommender System in Tourist Environment

Guillermo González-Suárez^I, Tatiana Delgado-Fernández^{II}, José Luis Capote-Fernández^I, Rafael Cruz-Iglesias^I

^I Empresa GeoSÍ, Grupo Empresarial GEOCUBA. La Habana, Cuba.

E-mail: guille@geomix.geocuba.cu , capote@geomix.geocuba.cu , rcruz@geomix.geocuba.cu

^{II} Universidad Tecnológica de La Habana. Facultad de Ingeniería Industria. La Habana, Cuba

E-mail: tdelgado@ind.cujae.edu.cu

Recibido: 25 de febrero de 2014

Aprobado: 8 de septiembre de 2016

RESUMEN

Este trabajo se centra en el desarrollo de un método de sistemas de recomendaciones híbrido sensible al contexto basado en ontologías en el entorno de las Infraestructuras de Datos Espaciales, que impacte positivamente en la efectividad de la toma de decisiones en las aplicaciones para usuarios de sistemas móviles en el sector turístico. El método elaborado cuenta de tres fases: marco semántico y de datos, otra que implementa el sistema de recomendaciones y una encargada de la visualización de las recomendaciones. La fase de sistema de recomendaciones implementa un filtro espacial, un filtro semántico y otro colaborativo. Las recomendaciones generadas muestran los elementos de posible interés, estas pueden ser visualizadas en forma tabular o combinadas con información espacial. Se comprobó por medio de un experimento y diferentes métricas aplicadas, el incremento en la efectividad del mismo para usuarios móviles según sus preferencias en cuanto a las recomendaciones dadas en un entorno espacial.

Palabras clave: Sistemas de Recomendaciones, Contexto, Ontologías, Infraestructura de Datos Espaciales.

ABSTRACT

This paper focuses on developing a method of hybrid context-aware recommender systems, based on ontologies in the environment of the Spatial Data Infrastructure, which will positively impact the effectiveness of decision-making in the tourist applications for mobile users. The method developed has three components: data and semantic framework, other that implements the recommender

system and another responsible for displaying the recommendations. The recommender system component implements a spatial, a semantic and collaborative filters. The recommendations generated show the elements of potential interest, these can be displayed in tabular or combined with spatial information. It was found through an experiment and different applied metrics the increase in its effectiveness for mobile users according to their preferences as to the recommendations given in a spatial environment.

Key words: Recommender System, Context, Ontology, Spatial Data Infrastructure.

I. INTRODUCCIÓN

El turismo representa actualmente un importante papel en la economía de la mayoría de los países desarrollados, estando fuertemente influenciado por las innovaciones tecnológicas [1]. Las empresas y los destinos turísticos deben ofrecer nuevos instrumentos y herramientas [2], y 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 [3]. 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 [4, 5]. Las TIC han cambiado la forma de difusión de la información turística de los destinos, además de la forma en la 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 [6].

El sector turístico en Cuba se ha convertido progresivamente en uno de los principales motores de la economía. Con 2 852 572 de turistas recibidos en el 2013 [7] y tres millones de turistas esperados en el año 2014, una parte cada vez mayor de ellos, prefieren rentar autos y viajar por los principales atractivos del país, no solo en los denominados "polos turísticos", sino por toda la Isla para conocer las costumbres y modo de vida de los cubanos. En sus travesías, no disponen de mucha información de los territorios que visitan, de forma tal que puedan ver satisfechas sus necesidades en relación a puntos de interés de acuerdo a sus propias preferencias.

Los destinos deben tener una visión amplia y dinámica y proporcionar al turista herramientas que ofrezcan información personalizada, tratando de integrar las diferentes necesidades existentes [8]. 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 [5].

Tanto en el sector turístico, como en la sociedad en general, existe una creciente necesidad de compartir información entre las diferentes partes interesadas y los diversos sistemas de información a fin de permitir un uso coherente y contextual de la misma.

Esta necesidad es la base para el surgimiento de una serie de infraestructuras a diferentes escalas, internacionales, nacionales y regionales para la recopilación y difusión de datos geográficos, las Infraestructuras de Datos Espaciales [9].

Una Infraestructura de Datos Espaciales (IDE en lo adelante) abarca las políticas, tecnologías, estándares y recursos humanos necesarios para la efectiva recolección, administración, acceso, entrega y utilización de los datos espaciales a diferentes niveles, en función de la toma de decisiones económicas, políticas y sociales y del desarrollo sostenible [10].

En la sociedad moderna existe un número creciente de información que está considerándose crítica para la toma de decisiones diarias. Como cada vez se hace asequible más cantidad de información en línea, que incluye algún contexto geográfico, la habilidad para descubrir y acceder a los recursos de datos geográficos para el uso en la visualización, planificación y toma de decisiones, se convierte en un requisito indispensable para apoyar a la sociedad [11].

Ante la gran cantidad de información existente en la Web y en algunos servicios que se proporcionan en la misma, en los últimos años se han ido desarrollando e implantando diferentes herramientas para facilitar a los usuarios un acceso rápido y adecuado a la información que necesitan [12].

En los últimos años, el uso de sistemas de recomendación se ha popularizado en aplicaciones de diverso tipo (*iTunes Store, Amazon, etc.*). En el ámbito del turismo, resulta de especial interés 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. Este hecho provoca que en muchos casos los usuarios se sientan desbordados al tener que elegir entre una gran diversidad de alternativas [13]. Por tanto, los sistemas de recomendaciones 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. Deben tomar en cuenta las preferencias de los usuarios para sugerir información de posible interés para ellos, ayudándolos a limitar sus búsquedas al brindar una serie de elementos que pudieran ser pertinentes. Proveer alertas de información y apoyar la navegación cuando se cuenta con grandes volúmenes de información [12]. En la literatura y en el mercado existen diversos tipos de sistemas de recomendaciones que difieren en el método o proceso de obtención de las recomendaciones y/o en las fuentes de información usadas, pudiendo destacar: sistemas de recomendaciones colaborativos [14, 15, 16], sistemas basados en contenido [17, 18], en conocimiento [19], en utilidad, o hibridando alguna de estas técnicas [20].

Actualmente el uso de aplicaciones móviles está constituyendo una auténtica revolución en el mercado de los *smartphones*, suponiendo un estímulo para el desarrollo de nuevos sistemas [13]. Por tanto, las aplicaciones móviles orientadas al turismo pueden convertirse en un nuevo canal que posibilite información, promoción y comercialización turística.

La revisión de trabajos relacionados evidencia diversas iniciativas en la definición y el diseño de modelos de contexto. Algunos de estos trabajos recogen interesantes propuestas teóricas, aunque no siempre se ha llevado a cabo su implementación. En este sentido en [21] se presentan destacadas tendencias de modelado, mientras que en [22, 23] se han propuesto diversos modelos híbridos de contexto. Sin embargo, en estas propuestas de modelo hay limitados acercamientos de modelos de contexto que combinen la aproximación basada en ontologías y la espacial [24].

De cara a su utilización en sistemas turísticos, una carencia importante detectada en estos modelos es que no utilizan atributos de contexto [25]. Para mejorar la calidad de las recomendaciones, el sistema debe también tener en cuenta por tanto, parámetros de contexto [26], tales como horario, clima, compañía, etc.

No se encontraron en la literatura consultada referencias de sistemas de recomendaciones para el turismo sensibles al contexto y basados en ontologías que agreguen valor a las IDE, vistas estas como los principales repositorios de información geográfica nacional generada por agencias de cartografías oficiales u otros proveedores con reputación, y validadas por su uso en diversos sectores de la sociedad.

Este trabajo describe un método de Sistemas de Recomendaciones para turistas (usuarios móviles) que, mediante la consulta a ontologías de puntos de interés que se encuentra en su contexto espacial, devuelven en cada momento aquellos potenciales lugares que pueden ser interesantes para el turista de acuerdo a su perfil y la experiencia anterior de usuarios "similares". Su viabilidad es demostrada a partir de la implementación del método en una aplicación de valor agregado a los servicios de la Infraestructura de Datos Espaciales de la República de Cuba.

El efecto de su uso para el sector del turismo es comentado en términos cualitativos, en tanto los beneficios potenciales para el mejoramiento de los servicios que ofrece el sector, y cuantitativamente, mediante estimados de incremento de valores al incorporar este tipo de servicio de cara al turista.

II. MÉTODOS

La información que nutre el sistema está representada mediante una ontología de destinos y una base de datos de usuarios donde se almacenan su perfil y preferencias. Se presenta el resultado del análisis y diseño del método propuesto, partiendo de lo general a lo particular, mostrando primero un esquema global, como parte de la descripción teórica del método, para posteriormente detallar los componentes participantes.

Tecnologías Utilizadas

Para el desarrollo del sistema propuesto por sus características como gestor de base de datos de código abierto se decidió utilizar *Postgre SQL* 9.2. La implementación de las funcionalidades de la

MÉTODO PARA SISTEMAS DE RECOMENDACIONES SENSIBLES AL CONTEXTO EN ENTORNOS TURÍSTICOS

aplicación se utilizó una serie de tecnologías, siempre priorizando las de software libre, que se enumeran a continuación. El lenguaje de programación seleccionado fue Java, utilizando como Interfaz de Desarrollo *NetBeans*. Se utilizó como marco de desarrollo (*framework*) *Google Web Toolkit* (GWT), creado por *Google*, permite ocultar la complejidad de varios aspectos de la tecnología *AJAX*. Es compatible con varios navegadores, lo cual es notorio, ya que cada navegador suele necesitar código específico para lograr un *front-end* correcto en una aplicación web.

Los ambientes visuales se lograron utilizando *Ext*, una biblioteca de clases *Open Source* en lenguaje *JavaScript*, que puede ser incluida en las aplicaciones Web e incluye una gran cantidad de componentes con una interfaz visual muy bien concebida. Para la visualización de los datos geográficos en el cliente se utilizó *Open Layers*, una biblioteca de código abierto, completamente libre, desarrollado en *JavaScript* y bajo licencia de distribución de software de *Berkeley* (BSD por sus siglas en inglés), su función es facilitar la ubicación de mapas dinámicos en una página Web. La calidad de las recomendaciones del método propuesto depende del componente que las realiza y lógicamente influyen en la eficiencia de la ejecución del mismo. Dentro de los motores existentes se escogió *Mahout*, una biblioteca de código abierto de aprendizaje automatizado del proyecto *Apache*.

II. RESULTADOS

El método creado está dividido en tres fases, Marco Semántico y de Datos, Sistema de Recomendaciones Sensible al Contexto Basado en Ontologías y Visualización de las Recomendaciones; la información fluye por él en ese orden. La figura 1 muestra el diagrama que sirve de descripción al método creado:

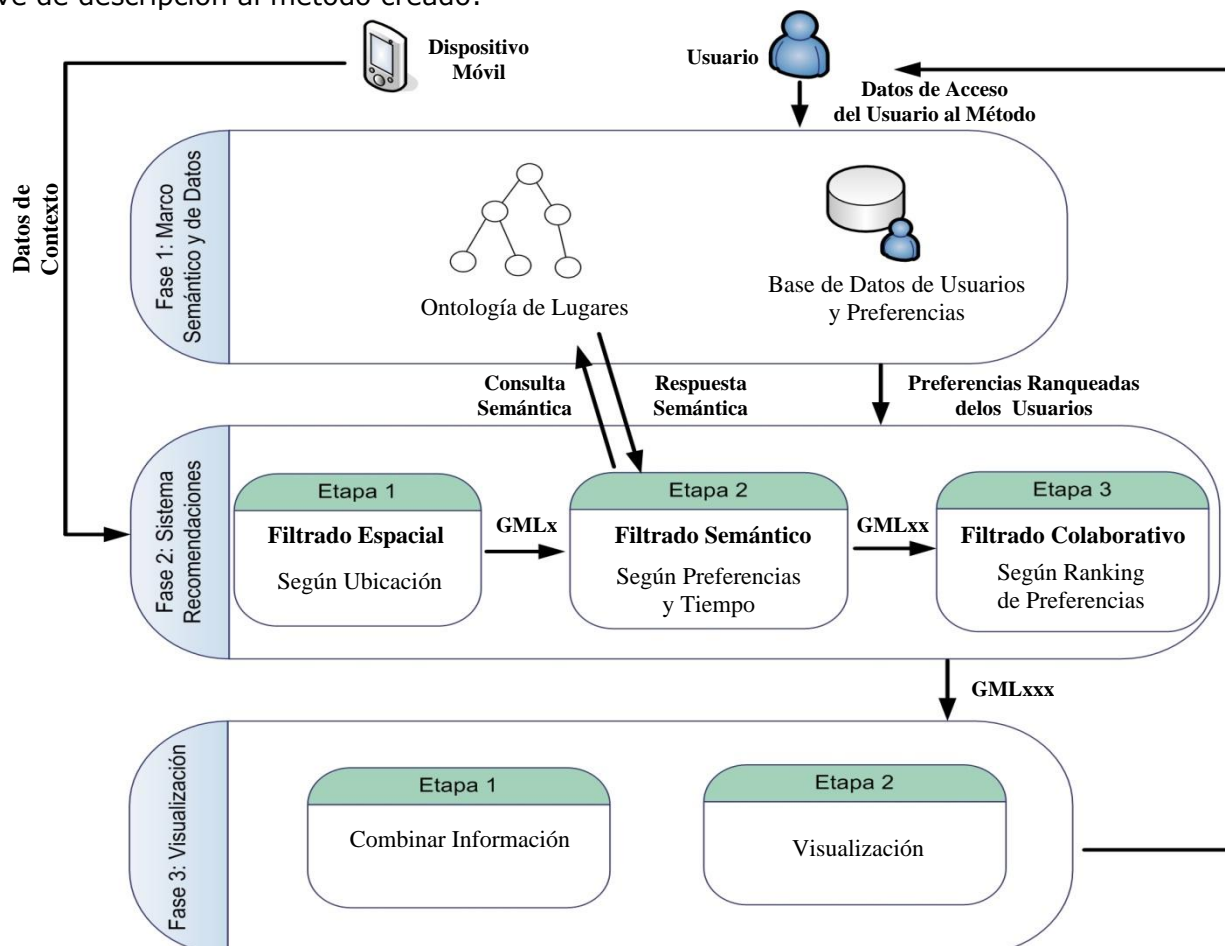


Fig. 1. Diagrama del Método de un Sistema de Recomendaciones Sensible al Contexto Basado en Ontologías

Su desempeño comienza cuando un usuario se registra en el sistema, en ese momento a partir de los datos de acceso se gestionan sus preferencias en la base de datos de Perfiles de Usuario. Estas se ordenan descendientemente y de conjunto con los datos de contexto. Los datos de su posición geográfica y de tiempo, constituyen la información de partida de la fase de sistema de recomendaciones, una fase clave dentro del método, encargada de filtrar los datos de los posibles destinos de este usuario mediante tres criterios, Filtrado Espacial, Semántico y Colaborativo. Por último, se ejecuta la fase de visualización del método, que tiene la función de combinar los datos obtenidos en la fase anterior con los servicios de mapas y obtener una representación lista para su visualización. La información geográfica dentro del método se maneja en formato GML [27]. A continuación, se describen en detalle cada una de las fases con sus elementos más importantes.

Fase 1: Marco Semántico y de Datos

Esta fase se inicia cuando el usuario (en este caso el turista) se registra en el sistema o cuando ocurre una modificación de su contexto, como puede ser un cambio de localización. Al final de su ejecución se obtienen las preferencias del usuario ordenadas por nivel de preferencia a partir de consultas a su base de datos, información que sirve de inicio a la fase de sistema de recomendaciones del método.

Como soporte de la información en el método propuesto se escogieron una ontología de Lugares y una base de datos de Perfiles de Usuario. La unión de ambas forma el marco semántico y de datos del método.

El marco ontológico que sirve de soporte al método está formado por una ontología de Lugares, que son los recursos con ubicaciones específicas que tienen atractivo para los perfiles de usuario, esta ha sido diseñada utilizando la herramienta *Protegé*, un entorno de código abierto de desarrollo de sistemas basados en conocimiento, uno de los editores de ontologías OWL más utilizados.

El objetivo de la ontología es soportar toda la información de los posibles destinos turísticos a visitar por los usuarios, para su posterior consulta semántica. Para su diseño, se partió de una super clase que se denomina Lugar, en la que se incluyen las propiedades de Nombre, Dirección, Teléfono, Latitud y Longitud, Horario, entre otras. A partir de ellas se generaron clases que heredan y se especializan en el lugar que van a almacenar, como por ejemplo, restaurantes, hoteles, cafeterías, playas, museos, etc.

Los sistemas de recomendaciones trabajan partiendo de la información que se conoce sobre el usuario, habitualmente almacenada en su perfil [28]. Como parte del método se incluye un proceso inicial donde se solicitan al usuario una serie de datos (nombre, edad, sexo, gustos generales, etc.) que contribuyen a inicializar sus perfiles. Además de estos, están registrados los datos de sus preferencias, en cuanto a lugares más visitados, esta información puede ser suministrada por el cliente o se registra en la medida en que se usa el sistema. El perfil del usuario es almacenado en una base de datos, lo que completa la fase inicial del método.

Fase 2: Sistema de Recomendaciones Sensible al Contexto

La fase 2 se nutre de la información proporcionada por la fase anterior. Para su inicio es indispensable que el usuario esté registrado y que se hayan calculado sus preferencias. Otro requisito importante es que el dispositivo móvil tenga información de su ubicación.

Durante su ejecución, que está dividida en tres etapas, se logra refinar el conjunto de posibles destinos turísticos a visitar por el usuario hasta llegar a los que más se adaptan a sus preferencias.

En la etapa uno, se hace un filtrado espacial de estos destinos, siguiendo el criterio de distancia espacial. En la segunda etapa se ejecuta un filtrado semántico, utilizando como criterio de búsqueda las preferencias de los usuarios, mientras que en la tercera y última etapa son procesados por un motor de recomendaciones que es capaz de ordenarlos en función de las preferencias y los gustos del usuario que los está consultando. En todas las etapas de esta fase los datos están en Lenguaje de Mercado Geográfico (GML).

Para la ejecución de la etapa de filtrado espacial es indispensable contar con los datos de posición del usuario, que son capturados por el dispositivo móvil. Esta etapa reduce considerablemente el universo de búsqueda del método, ya que se eliminan aquellos destinos que no están dentro del área de influencia. Los resultados devueltos se almacenan en un fichero en formato GML, que sirve de

MÉTODO PARA SISTEMAS DE RECOMENDACIONES SENSIBLES AL CONTEXTO EN ENTORNOS TURÍSTICOS

fuentes de información para la siguiente etapa. Para la selección de los destinos incluidos en el área de influencia se elaboran consultas semánticas en formato GeoSPARQL, como se muestra a continuación:

```
PREFIX geo: http://www.opengis.net/ont/OGC-GeoSPARQL/1
PREFIX geof: http://www.opengis.net/def/queryLanguage/OGC-GeoSPARQL
SELECT ?lugar WHERE {?lugar geo:hasGeometry ?pgeo.
FILTER (geof:distance (?pgeo, "PONT((-80.089005 23.913574))"^^geo-sf:WKTLiteral), units:m) < 2000)}
```

La consulta anterior está elaborada para solicitar los lugares (lugar) pertenecientes a la ontología de destinos turísticos que se encuentran a menos de 2000 metros (units: m < 2000) del sitio donde se encuentra el usuario que se corresponde con las coordenadas: (-80.089005 23.913574).

Con el fin de seleccionar un elemento p, dentro de los puntos extraídos de la etapa anterior, en la etapa de Filtrado Semántico se utiliza el filtro semántico implementado en la IDERC[29], que permite realizar una selección de los lugares para representar la clasificación de todas las actividades vinculadas al contexto en P y el administrador de perfiles para obtener los datos de los usuarios en U. Esta solicitud es traducida y convertida a formato de filtro semántico, que el componente relaciona con los criterios solicitados al administrador de contexto de usuario y como resultado de la ejecución se obtiene una lista de lugares que coinciden con el contexto. A continuación se muestra un ejemplo de esta consulta:

```
PREFIX geo: http://www.opengis.net/ont/OGC-GeoSPARQL/1
PREFIX geof: http://www.opengis.net/def/queryLanguage/OGC-GeoSPARQL/
SELECT ?lugar WHERE{?lugar a place: Restaurante }
```

En esta consulta, expresada de forma sencilla y abreviada para la comprensión del lector, se están seleccionando de la ontología de destinos turísticos, los lugares que en su descripción coinciden con Restaurante. Como la semántica de estos destinos está descrita por una ontología, el motor de búsqueda semántica es consciente de la jerarquía de clases de cada uno de ellos. Esto significa que el motor puede seleccionar la estrategia de predicción apropiada para cada clase de destino.

En la última etapa de esta fase, denominada Etapa de Filtrado Colaborativo, es fundamental un elemento, el motor de recomendación, que utiliza múltiples estrategias para predecir cómo cada destino turístico puede responder a las preferencias del usuario. Una estrategia selecciona y/o combina múltiples técnicas de predicción para decidir cuál es la más adecuada para proporcionar una recomendación basada en la información más actualizada del conocimiento del usuario.

El motor de recomendaciones utiliza una base de datos de elementos y usuarios para generar predicciones. Primeramente emplea técnicas estadísticas para encontrar a vecinos, es decir usuarios con un historial de valoraciones sobre los elementos similares al usuario actual. Una vez que se ha construido una lista de vecinos se combinan sus preferencias para generar una lista con los N elementos más recomendables para el usuario actual. Esta técnica de recomendación se denomina "vecinos más cercanos".

En primer lugar es necesario medir los parecidos de todos los usuarios con el usuario actual, para ello se ejecutan funciones que permiten calcular su grado de similitud y vecindad. En el caso del método se utilizan varios métodos de calcular las similitudes uno de ellos es la función de correlación de *Pearson* definida como sigue: ver ecuación 1

$$w_{a,u} = \frac{\sum_{i=1}^m [(r_{a,i} - \bar{r}_a)(r_{u,i} - \bar{r}_u)]}{\sqrt{\sum_{i=1}^m (r_{a,i} - \bar{r}_a)^2 \sum_{i=1}^m (r_{u,i} - \bar{r}_u)^2}} \quad (1)$$

donde:

$w_{a,u}$ es el valor (peso) de la similitud entre el usuario activo a y su vecino u .
 m es el número de elementos.

$r_{u,i}$ es el valor de preferencia asignado por el usuario u al elemento i .

\bar{r}_u representa la media de todos los valores asignados por el usuario u .

Después de compiladas todas las similitudes entre los usuarios, la vecindad está formada y el motor está en condiciones de realizar las recomendaciones. Para ello, el método se sigue por la siguiente función de predicción: ver ecuación 2

$$P_{a,i} = \bar{r}_a + \frac{\sum_{u=1}^n [(r_{u,i} - \bar{r}_u) * w_{a,u}]}{\sum_{u=1}^n w_{a,u}} \quad (2)$$

donde:

$w_{a,u}$ es el valor (peso) de la similitud entre el usuario activo a y su vecino u .

$P_{a,i}$ representa la predicción para el usuario activo (a) respecto a elemento i .

n es el número de vecinos.

$r_{u,i}$ es el valor de preferencia asignado por el usuario u al elemento i .

\bar{r}_u representa la media de todos los valores asignados por el usuario u .

Una vez que el motor de recomendaciones termina su ejecución, se logra una lista ordenada de lugares turísticos, donde los primeros son los que más interesan al usuario, estos puntos son incorporados a un fichero en formato GML y están listos para pasar a la tercera y última fase.

Fase 3: Visualización de las Recomendaciones

Esta fase se ejecuta cuando ya se ha realizado la selección de los posibles destinos que se recomendarán al usuario. Como punto de partida se tiene un fichero GML que se obtiene de la etapa de Filtrado Colaborativo de la fase anterior. Esta tercera fase cuenta con dos etapas.

La información resultante de la fase Sistema de Recomendaciones viene codificada en un fichero de formato GML, que contiene un listado de lugares sensibles de ser recomendados al usuario. En la etapa de combinación de la información se propone mezclar estos datos con información geográfica proveniente de servicios WMS [30] o WFS [31], que son especificaciones estandarizadas de servicios de mapas o de elementos geográficos provenientes del consorcio OGC [32], y obtener la representación de estos puntos sobre una cartografía de referencia.

Por medio de estos servicios se pueden solicitar datos de servicios de información geográfica ubicados en cualquier lugar del mundo, siguiendo las especificaciones dictadas para ello. En esta etapa del método, es posible combinar la información obtenida con servicios como *Google Earth*TM, *Yahoo Maps* y/o los servicios publicados en una Infraestructura de Datos Espaciales, como la IDERC (www.iderc.co.cu).

El resultado de esta integración de servicios e información se traduce en una o varias consultas a servicios de datos espaciales y los datos en formato GML, que serán entregados a un componente de visualización, elementos que sirven de base a la etapa de Visualización de la información, última del método.

En esta etapa se proponen dos formas de visualización del resultado del método, la primera consiste en visualizar los datos en una ventana temática o sea visualizarlos en un formulario donde se puedan apreciar varias de sus propiedades. La otra, es transferir el fichero con la información a un componente de visualización espacial (*Open Layers*) para que pueda ser combinado con otra información de infraestructura o de imágenes satelitales que sirva de base a la visualización de los datos.

El método descrito define un flujo de trabajo para la ejecución de un sistema de recomendaciones sensibles al contexto basado en ontologías para usuarios de dispositivos móviles en entornos de IDE. Orquesta las diferentes operaciones a realizar ejecutando filtros espaciales, semánticos y colaborativos de manera lógica. Incluye un pre-filtrado contextual que garantiza la vinculación al entorno del usuario y la utilización de implementaciones de filtrado colaborativos en formato 2D.

MÉTODO PARA SISTEMAS DE RECOMENDACIONES SENSIBLES AL CONTEXTO EN ENTORNOS TURÍSTICOS

Los datos obtenidos como resultado de la Fase de Sistema de Recomendaciones Sensible al Contexto del método, se mezclan sin dificultad con el servicio de mapas de la IDERC, lo que garantiza su estandarización, pudiendo en otros escenarios de uso vincularlos a otros servicios de mapas que cumplan con los estándares OGC.

La validación del método propuesto se centró en la fase de sistema de recomendaciones sensible al contexto basado en ontologías (Fase 2). Un elemento esencial en la validación es contar con los datos necesarios y representativos para la realización de la misma.

Como principal problema para realizar las pruebas se encontró que el sistema no contaba con perfiles de usuarios reales (datos personales, preferencias generales, etc.). Para obtener el perfil de usuario, resultó conveniente simular la utilización del sistema por parte del usuario, realizando encuestas a usuarios reales.

Concretamente, se realizaron encuestas a 60 usuarios diferentes, 36 hombres y 24 mujeres. Los encuestados están distribuidos en varios rangos de edades y nivel educacional, como está resumido en tabla 1, los voluntarios fueron seleccionados de forma tal que no hubiese diferencias significativas en su distribución demográfica y estuvieran representados la mayor cantidad de individuos posibles.

Tabla 1. Información demográfica extraída de la encuesta

Dimensión demográfica	Cantidad	% del total
Género		
Femenino	24	40
Masculino	36	60
Rango de Edades		
Entre 13 y 18 años	7	11,66
Entre 19 y 24 años	21	35
Entre 25 y 50 años	24	40
Mayor de 50 años	8	13,33
Nivel Educativo		
Medio	2	3,33
Medio Superior	23	38,33
Superior	20	33,33
Otro	15	25
Ocupación		
Estudiante	15	25
Obrero	13	21,66
Profesional	20	33,33
Ama de casa	7	11,66
Jubilado	5	8,33

Además se cuenta con una ontología de destinos turísticos, en la que existe información de 2 946 destinos en Cuba, distribuidos en varias clasificaciones entre las que están restaurantes, cafeterías, hospitales, farmacias, centros nocturnos y varias más. De cada punto se tiene además de la ubicación, la categoría, tipo de servicio y horario, entre otros datos, que permiten vincularlos con las preferencias de los usuarios.

A partir de la combinación de esta información recopilada se pudo obtener un universo de pruebas que permitió establecer las preferencias de cada usuario por cada uno de los elementos de la ontología de lugares. Este universo de datos estaría formado por 176 760 lugares con su escala de preferencia.

El proceso de validación de esta etapa consiste en conocer si las recomendaciones se ajustan a los gustos del usuario, es por ello que las pruebas de esta etapa cobran especial importancia. Para que los datos de evaluación sean fiables, se deben realizar múltiples ejecuciones y calcular la media de las métricas de evaluación. Esto disminuye la probabilidad de que los resultados del algoritmo sean mejores o peores debido a los datos que se hayan elegido para entrenamiento y evaluación.

Para este tipo de validación se aplican protocolos de evaluación, estos determinan qué datos conoce un algoritmo durante la construcción del modelo (fase de entrenamiento) y qué datos debe calcular a partir del mismo (fase de evaluación). Con estas predicciones se calcula si el algoritmo es eficaz y en qué medida [33].

Una configuración ampliamente utilizada en la evaluación de algoritmos de sistemas de recomendaciones es 80 % entrenamiento, 20% evaluación. Esto quiere decir que el 80% de los datos iniciales se utilizarán para la generación del modelo y el 20 % restante se utilizará para comprobar la eficacia del mismo [34].

Para definir qué tan buenas son las recomendaciones generadas se utilizan las métricas de evaluación, que miden la calidad de los resultados, con el objetivo de analizar los puntos fuertes y débiles, visualizar las variables involucradas y ver cómo afectan los diferentes parámetros y sus variaciones.

Para demostrar la dependencia existente entre el conjunto de entrenamiento y la precisión de las recomendaciones, se dividió este en 8 subconjuntos de forma aleatoria, tomado para ello a partir del 30 % de los datos hasta llegar al conjunto total de entrenamiento; con cada uno de ellos se entrenó el método, se ejecutaron las recomendaciones y se calculó el índice de error medio absoluto, utilizando la expresión 3.

$$MAE = \frac{\sum_{i=1}^N |p_i - r_i|}{N} \quad (3)$$

donde:

p_i : es el valor de la recomendación calculada por el método.

r_i : es el valor que el usuario ha expresado de su preferencia por el elemento i .

N : es la cantidad de elementos del conjunto.

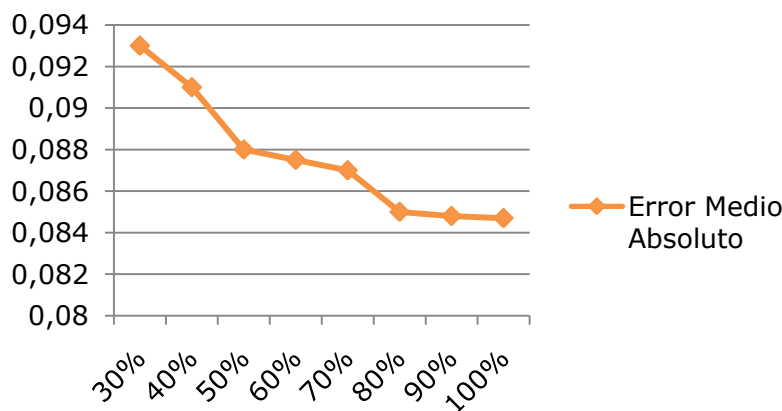


Fig. 2. Comportamiento del Error Medio Absoluto cambiando el porcentaje de aprendizaje

Para realizar las pruebas de precisión se generaron un promedio de 10 recomendaciones por usuario. Las métricas utilizadas fueron Precisión y Recuerdo, para su funcionamiento necesitan que se defina un criterio de relevancia, que se utiliza para decidir si el producto es relevante o no. En este caso, recordando que los usuarios establecen sus preferencias en una escala del 1 al 10, se considera que los productos con una valoración mayor o igual a 7 son relevantes y menor o igual a 6 no lo son.

El proceso es el siguiente: se simula el registro de cada usuario asociado al conjunto de prueba en el sistema, ingresando la mínima información conocida del usuario (datos personales). Seguidamente el sistema genera una recomendación, una selección S para el usuario. Con la intersección entre S y los lugares que realmente ha visitado el usuario, se denomina R (obtenidos en el cuestionario), se obtiene el conjunto de lugares relevantes y seleccionados.

MÉTODO PARA SISTEMAS DE RECOMENDACIONES SENSIBLES AL CONTEXTO EN ENTORNOS TURÍSTICOS

Las pruebas de precisión se realizaron utilizando la expresión 4:

$$\text{Precisión} = \frac{N_{rs}}{N_s} \quad (4)$$

donde:

N_{rs} es el número de elementos relevantes seleccionados por el sistema.

N_s es el número total de elementos seleccionados por el sistema.

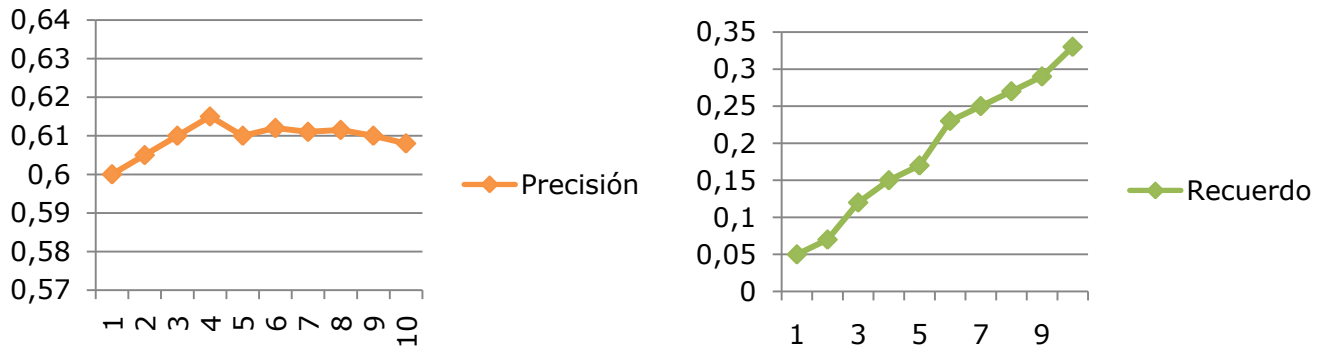


Fig. 3. Precisión y recuerdo de los algoritmos en el experimento

La evaluación del recuerdo se realizó a partir de la expresión 5:

$$\text{Recuerdo} = \frac{N_{rs}}{N_r} \quad (5)$$

donde:

N_r es el número de elementos que el usuario ha clasificado como relevantes.

N_{rs} es el número de elementos relevantes seleccionados por el sistema.

IV. DISCUSIÓN

Como se puede ver en las figuras 2 y 3 del acápite anterior, cuanto más aumenta el número de elementos que sirven de entrenamiento al método, se reduce el Error Medio Absoluto, por lo que se puede afirmar que el sistema de recomendaciones propuesto mejora la calidad de sus predicciones. Aunque cuando el número de datos que sirve de entrenamiento es muy elevado (para mayores de un 80 %) se observa que existe una tendencia a la linealidad. Esto puede ser debido a la dificultad de encontrar puntos de interés que aún no han sido visitados por el usuario y que cumplan sus expectativas.

El comportamiento de la precisión durante la realización de las pruebas, se puede observar que el algoritmo diseñado gana en precisión para 4 o menos recomendaciones, lo que indica cuál es su mayor capacidad de recomendar destinos correctamente. Por lo tanto, se puede concluir que el algoritmo es útil para ser utilizado en aquellos dominios en los que el usuario reciba una lista de recomendaciones pequeña, por ejemplo, cuando el sistema ofrece recomendaciones al usuario sin que estas hayan sido solicitadas previamente.

El algoritmo propuesto mejora en Recuerdo en la medida que aumenta la cantidad de recomendaciones, esto indica que la ponderación supone una mejora; lo que confirma que los usuarios, al realizar las valoraciones de preferencia, tienen en cuenta las características de los puntos de interés.

Al evaluar los resultados del algoritmo de Filtrado Colaborativo mediante las métricas de Error Medio Absoluto, Precisión y Recuerdo se demostró que las predicciones calculadas están en el rango de las puntuaciones reales dadas por los usuarios.

En el caso de estudio se demuestra la validez del modelo propuesto, revelando que, para un número de recomendaciones menor o igual a 4, obtiene mejores resultados que para mayor cantidad de recomendaciones. Esta condición hace al algoritmo óptimo para ser utilizado en aplicaciones móviles, donde el usuario no solicita las recomendaciones, sino que estas se realizan de forma automática.

Con la obtención de este método se logra un incremento en las capacidades analíticas de los usuarios de sistemas móviles, a partir de su aplicación se aumentó la efectividad de la toma de decisiones de los clientes de la industria del turismo en nuestro país y se elevó la capacidad de análisis en las aplicaciones de cara a los turistas. Su implementación mejora la calidad de la información obtenida por los turistas que hacen uso de los servicios de la IDERC y simplifica los resultados que llegan a cada uno de los dispositivos desde los que se accede.

El efecto económico de la aplicación del método obtenido se realizó en una implementación del mismo, que sirve como escenario actual de comercialización, específicamente en los autos de alquiler de lujo para el sector turístico. El costo total de propiedad del sistema implementado para el sector del turismo es de \$ 64 420,56. De este total el valor de la inversión asciende a \$ 32 856,49 para un 51%, siendo el restante de \$ 31 564,07, el 49 %, el valor de operaciones para el primer año. A partir del modelo de comercialización estudiado se estableció un precio por uso del sistema desarrollado de \$2,00 por encima del precio del alquiler del auto. Con este valor de servicio, a partir del alquiler de 90 vehículos con el sistema, que representa el aproximadamente el 6,49 % de los rentados en un mes, se recupera el valor de la inversión en la implementación.

V. CONCLUSIONES

1. Los dispositivos móviles se han convertido en una tecnología cotidiana, estos han evolucionado a un nuevo paradigma, la sensibilidad al contexto. Las herramientas que filtran la información que reciben los usuarios de estos dispositivos han emergido como un elemento muy útil, dentro de ellas los sistemas de recomendaciones, ocupan un papel protagónico.
2. Existen varios modelos y métodos que combinan las tecnologías de los sistemas de recomendaciones, la sensibilidad al contexto y las ontologías dos a dos, actualmente implementados para dispositivos móviles. Pero no existe un método que combine las tres tecnologías en un solo proceso, desaprovechando las ventajas que esto puede representar en la recuperación y personalización de la información.
3. La viabilidad del método propuesto quedó demostrada con la implementación de un prototipo funcional que integró todos los componentes descritos en él. Complementariamente se implementó una aplicación Web que, a partir de simulaciones, permitió determinar su eficacia.
4. El método es efectivo en la recuperación de información por usuarios móviles, quedando demostrado en la implementación del método y la evaluación de las métricas (Error Medio Absoluto, Precisión y Recuerdo) del Sistema de Recomendaciones generado en el experimento.
5. Para el sector turístico, tanto desde el punto de vista cualitativo, mejorando la calidad de los servicios que se ofrecen, como desde una perspectiva económica, mediante una estimación económica de la habilitación del Sistema de Recomendaciones gradualmente, se puede percibir que este método es eficaz y eficiente. 📄

VI. REFERENCIAS

1. Sundbo J, Orfila S, Sorensen F. The innovative behaviour of tourism firms. Comparative studies of Denmark and Spain. Research Policy. 2007;36(1):88-106. ISSN 0048-7333.
2. Buhalis D, Law R. Progress in information technology and tourism management: 20 years on and 10 years after the Internet. The state of eTourism research. Tourism Management. 2008;29(4):609-23. ISSN 0261-5177.
3. Castejón R, Méndez E. Introducción a la economía para turismo: Editorial Prentice-Hall; 2012. ISBN 978-84-8322-522-6.
4. Juaneda CN, Riera A. La oportunidad de la investigación en economía del turismo. Estudios de Economía Aplicada. 2011;29(3):711-22. ISSN 1697-5731

MÉTODO PARA SISTEMAS DE RECOMENDACIONES SENSIBLES AL CONTEXTO EN ENTORNOS TURÍSTICOS

5. Guevara A, Aguayo A, Gómez I, et al. Sistemas informáticos aplicados al turismo: Editorial Pirámide; 2009. ISBN 978-84-3682-314-1.
6. Xiang Z, Pan B. Travel queries on cities in the United States: Implications for search engine marketing for tourist destinations. *Tourism Management*. 2011;32(1):88-97. ISSN 0261-5177.
7. ONEI. Anuario Estadístico de Cuba 2013. La Habana, Cuba: Dirección Nacional de Estadística; 2013.
8. Valdés L, Valle E, Sustacha I. El conocimiento del turismo en el ámbito regional. *Cuadernos de Turismo*. 2011 (27):931-52. ISSN 1989-4635.
9. Vaccari L, Shvaiko P, Marchese M. A geo-service semantic integration in Spatial Data Infrastructures International Journal of Spatial Data Infrastructures Research. 2009;4:24-51. ISSN. DOI 1725-0463.
10. Delgado Fernández T, Capote Fernández JL. Semántica Espacial y Descubrimiento de Conocimiento para Desarrollo Sostenible. En: Marco teórico de Infraestructuras de Datos Espaciales en el Proyecto CYTED IDEDES. 2009. p. 21-32.
11. Nebert Douglas D. Spatial Data Infrastructure Cookbook v2.0. 2004.
12. Espinilla M. SR-REJA. Sistema de Recomendación Híbrido.Georreferenciado [tesis de grado]. España2009.
13. Lymberopoulos D, Zhao P, König A, et al. Location-aware click prediction in mobile local search. 2011:413-22. ISSN 1557-735X
14. Adomavicius G, Tuzhilin A. Toward The Next Generation Of Recommender Systems: A Survey Of The State-Of-The-Art And Possible Extensions. *Knowledge and Data Engineering IEEE Transactions*. 2005;17(6). ISSN 1041-4347.
15. Sarwar B, Karypis G. Item-Based Collaborative Filtering Recommendation Algorithms. In: Proceedings of the 10th international conference on World Wide Web; Hong Kong; 2001.
16. Herlocker J, Konstan J. An Algorithmic Framework For Performing Collaborative Filtering. . In: The 22nd Annual International Acm Sigir Conference On Research And Development In Information Retrieval; Berkeley, USA: 1999.
17. Martínez L, Pérez LG, Barranco M. A Multigranular Linguistic Content-Based Recommendation Model. *International Journal Of Intelligent Systems* 2007;22(5). ISSN 1098-111X.
18. Pazzani M, Muramatsu J, Billsus D. Syskill Webert: Identifying Interesting Web Sites. *AAAI*. 1997;1:54-61. ISSN 0453-4662.
19. Burke R. Knowledge-Based Recommender Systems2000. ISBN 9780849397127. DOI [Citado
20. Hybrid. Recommender Systems: Survey and Experiments. User modeling and user-adapted interaction. 2002;12(4). ISSN 1573-1391.
21. Bettini C, Brdicza O, Henriksen K, et al. A survey of context modelling and reasoning techniques. *Pervasive and Mobile Computing*. 2010;6(2). ISSN 1574-1192.
22. Lee D, Meier R. Primary - Context Model and Ontology: A Combined Approach for Pervasive Transportation Services. En: Fifth Annual IEEE International Conference Pervasive Computing and Communications Workshops. Dublin, Ireland p. ISBN 978-0-7695-2787-1.
23. Park M, Sug Gu M, Ho Ryu K. Context information model using ontologies and rules based on Spatial Object. In: Advanced Intelligent Computing Theories and Applications With Aspects of Contemporary Intelligent Computing Techniques; Berlin, Hedelberg Springer; 2007.
24. Becker CH, Nicklas D. Where do spatial context models end and where do ontologies start? A proposal of a combined approach. In: Proceedings of the first international workshop on advanced context modelling, reasoning and management in conjunction with UbiComp; 2004.
25. Adomavicius G, Tuzhilin A, Berkovsky S, et al. Context-awareness in recommender systems: research workshop and movie recommendation challenge. *RecSys*. 2010:385-96. ISSN 1613-0073.
26. Lamsfus C, Grün C, Alzua Sorzabal A, et al. Crear vínculos basados en el contexto para mejorar las experiencias de los turistas. *Novática, Revista de la Asociación de Técnicos de Informática*. 2010 (203):17-23. ISSN 0211-2124.
27. Cox S, Daisey P, Lake R. OpenGIS Geography Markup Language (GML) Implementation Specification. OpenGIS project document reference number OGC. 2003.

28. Hong SK. Ubiquitous Geographic Information (UBGI) and address standards: ISO Workshop on address standards. En: Considering the issues related to an international address standard,. Copenhagen, Denmark. p. ISBN 978-1-86854-689-3.
29. Capote JL. Modelo de Servicios Semánticos en la IDERC [tesis doctoral]. La Habana, Cuba: Instituto Técnico Militar José Martí; 2011.
30. OGC. OGC Web Map Server Interface (WMS) Implementation Specification. OGC Document No. 00-028. 2000.
31. OGC. Web Feature Service (WFS) Implementation Specification; 2005.
32. Open GeoSpatial. Consortium. [Citado Disponible en: <http://www.opengeospatial.org>
33. Herlocker J, Konstan J, Terveen L, et al. Evaluating collaborative filtering recommender systems. ACM Transactions on Information Systems 2004;22(1):5-53.
34. Castro Gallardo J. Un nuevo modelo ponderado para Sistemas de Recomendación Basados en Contenido con medidas de contingencia y entropía [tesis doctoral]. España: Universidad de Jaen; 2012.