

Esbozo de una técnica para la recomendación de productos turísticos en Cuba usando GIS

Jesse Daniel Cano Otero¹, Marlon Jorge Remedios Gonzalez¹, Maria Teresa Morell Franco²

¹Universidad de las Ciencias Informáticas. Facultad I. Departamento de Programación.
{jdcano, mjremedios}@uci.cu

²Grupo Empresarial BIOCUBAFARMA. Departamento de Informática
maria@sd.biocubafarma.cu

Resumen. El turismo cubano se ha desarrollado tanto en los últimos 30 años al punto de convertirse en uno de los motores de la economía cubana. Con este desarrollo también han crecido las empresas cubanas que apuestan por el turismo electrónico como vía para dar a conocer sus productos y captar clientes. A pesar de este hecho, la mayoría de los sitios web cubanos de temática turística se limitan a brindar información de los distintos productos y servicios lo que provoca, en muchos casos, que el usuario se abrume con el cúmulo de información y abandone la búsqueda antes de encontrar un producto que se adecue a sus necesidades. La personalización ha sido reconocida como un factor crítico para las empresas de turismo electrónico exitosas y la utilización de sistemas de recomendación es el mejor enfoque para tratar con el problema de la personalización. El presente trabajo tiene como objetivo realizar un acercamiento preliminar a una técnica que permita obtener predicciones sobre a qué productos un usuario en particular daría una mejor valoración; estos productos serían los que se le mostrarían en primera instancia. Para lograr esto en el trabajo se analizan los elementos teóricos del entorno turístico en Cuba; de los sistemas de recomendación y de los sistemas de información geográfica como herramientas para la representación de la información. Por último se definen, por cada componente estructural identificado, un conjunto de reglas que permitirán obtener un sistema de turismo electrónico que maneje la personalización del servicio que presta de manera efectiva.

Palabras Clave. Productos turísticos; recomendación; sistema de información geográfica; técnica.

1 Introducción

Luego de la caída del campo socialista, el gobierno revolucionario apostó por el desarrollo turístico en Cuba como vía de proveer un alivio a la economía. Este sector se desarrolló al

punto de convertirse en uno de los motores de la economía cubana. Con el auge de internet en los noventa la industria turística mundial fue cambiando dando lugar al turismo electrónico (e-turismo). De hecho, varios estudios recientes indican que la Web se ha convertido en la principal fuente de información en el mundo occidental para las personas que buscan obtener información de los destinos turísticos cuando se encuentran planificando sus viajes [1]. A pesar de este hecho, la mayoría de los sitios web cubanos de temática turística se limitan a brindar información de los distintos destinos y en algunos casos brindan la posibilidad de realizar reservaciones para algún que otro producto turístico. Es obvio que la cantidad de información que recibe el usuario cuando desea encontrar un producto turístico cubano en la web puede convertirse en un problema porque el usuario se vería abrumado por el cúmulo de información disponible y es probable que abandone el sitio web antes de encontrar el producto de su preferencia.

La personalización ha sido reconocida como un factor crítico para las empresas de e-turismo exitosas [2]. El enfoque más exitoso para lidiar con el problema de la personalización es la utilización de los sistemas de recomendación (SR y SRs para el plural) [3,4]. Este enfoque provee a los usuarios técnicas para el filtrado de información usando el conocimiento sobre sus preferencias. Los SRs pueden proveer al visitante del sitio un conjunto de opciones, desde destinos, hoteles, museos y lugares de interés que se ajusten al perfil y presupuesto del visitante. Para la personalización del servicio en el e-turismo, los enfoques más populares son el uso de SRs con filtrado colaborativo y basados en el contenido.

Por otro lado, muchos SRs hacen uso de sistemas de información geográfica (GIS por sus siglas en inglés) como una vía de mejorar la presentación de las recomendaciones y, de esta manera, incrementar la interactividad del sistema con los usuarios. La utilización de GIS para el e-turismo es totalmente razonable dado a que la mayoría de la información turística tiene una estrecha relación con el espacio [5].

Dada la dinámica del sector turístico y el emergente, pero actualmente muy competitivo, mercado electrónico del turismo, casi todas las organizaciones presentes en la industria han implementado una estrategia que explota las tecnologías de la información (TI) e internet [6]. Esto sucede porque la utilización de las TI ha permitido que los turistas tomen un rol más activo en la producción de productos turísticos porque ya no se conforman con productos estandarizados [8].

Dependiendo de la localización de los clientes respecto al índice de inclusión digital [7], más y más clientes adquieren sus productos turísticos a través de sitios web y perciben que la imagen del sitio y su usabilidad afectan directamente sus intenciones adquisitivas [9]. Por lo tanto, entender las percepciones de los clientes y sus comportamientos online son determinantes para el desarrollo de un sitio de e-turismo exitoso [10]. Un SR puede ser de gran utilidad ya que provee asistencia en el proceso de brindar – recibir indicaciones acerca de las opciones que se ajustan mejor en un caso específico para individuos específicos [11]. Ricci [12] dice que un SR puede brindar información valiosa para asistir en el proceso de toma de decisiones del cliente a partir de la identificación de los

requisitos de este y mediante la identificación de su correlación con los requisitos y preferencias de otros clientes. Kaplanidou y Vogt [13] demostraron que la utilidad, para un usuario en particular, de un sitio web constituye un predictor significativo de la intención de viajar al destino.

Con la rapidez de las transmisiones de datos en internet, los tiempos de respuesta esperados de las organizaciones para con los usuarios han sido reducidos drásticamente. La reacción de los sistemas a las consultas online pueden, por lo tanto, influenciar la satisfacción del cliente y su comportamiento adquisitivo. Lograr un entendimiento de los distintos comportamientos online de los clientes puede incrementar la posibilidad de que el cliente complete la transacción [14]. Teniendo en cuenta esto, las características del comportamiento de compras online de los viajeros han sido estudiadas en profundidad por los investigadores. Por ejemplo, los estudios han mostrado que los clientes chinos son menos propensos a confiar en los anuncios publicitarios de productos turísticos y se enfocan comúnmente en el word-of-mouth (comentario, sugerencia, rumor) electrónico y en la información online [14].

A pesar de la gran importancia que, para Cuba, tiene la industria del turismo las empresas que utilizan el e-turismo como vía de captar clientes saturan de información la Web de manera que los potenciales clientes se enfrentan al problema de determinar qué producto turístico consumir a partir de un monto abrumador de información. Esto invariablemente atenta contra la eficiencia con que operan estas empresas y contra el consumo de los productos turísticos que se ofertan. A las organizaciones dedicadas al turismo que han fallado en aplicar correctamente las tecnologías de la información y las comunicaciones (TICs) les es muy difícil gestionar sus negocios sensibles a la información sin afectar su competitividad [15]. Esta situación puede ser mitigada con el uso de un método que permita predecir qué producto turístico puede ser de interés para un grupo determinado de clientes. En este contexto se puede hacer uso de un SR, pero, para esto, se deben caracterizar los distintos elementos que componen el entorno turístico cubano con el objetivo de proveer una recomendación de productos turísticos con exactitud y de manera precisa. En el presente trabajo se esbozará una técnica para obtener recomendaciones efectivas de productos turísticos en Cuba a partir de la definición de un conjunto de reglas que indicarán cómo deben ser implementados cada uno de sus componentes.

2 Sistemas de recomendación

Los SRs son herramientas de software y técnicas que proveen sugerencias o recomendaciones acerca de los ítems que pueden ser útiles para un usuario [4,16]. Desde este punto de vista se puede decir, entonces, que los SRs son la respuesta técnica al hecho de que las personas frecuentemente se basan en la experiencia y recomendaciones de otras personas cuando se enfrentan a un nuevo campo del conocimiento, cuando no conocen con amplitud todos los hechos o cuando el conocimiento excede la cantidad de información que un

humano puede manejar cognoscitivamente [8]. Estas recomendaciones se relacionan a varios procesos de tomas de decisiones, por ejemplo: qué ítem comprar, qué canción escuchar o qué noticia leer.

El término “ítem” se usa, de manera general, para denotar qué es recomendado a los usuarios por el sistema. Los SRs normalmente se centran en un tipo de ítem específico (dígase canciones, noticias o libros) y de acuerdo con esto su diseño, su interfaz gráfica de usuario (GUI por sus siglas en inglés), y la técnica de recomendación se configura para proveer sugerencias útiles y efectivas para ese tipo específico de ítem.

Luego, los SRs se dirigen principalmente a los individuos que carecen de la suficiente experiencia personal o la competencia para evaluar el potencialmente abrumador número de alternativas que un sitio Web, por ejemplo, puede ofrecer [4].

En su forma más simple, las recomendaciones se ofrecen como una lista ordenada de ítems. Al ejecutar este ordenamiento, los SRs tratan de predecir cuáles son los productos o servicios más convenientes para el usuario. Para lograr esto, computacionalmente, los SRs recolectan las preferencias de los usuarios, ya sea explícitamente, por ejemplo: solicitando a los usuarios valoraciones sobre los ítems, o infiriéndolas a partir de las acciones de los usuarios. En la actualidad, los SRs han demostrado ser un medio valioso de lidiar con el problema del exceso de información.

La taxonomía propuesta en el artículo “Hybrid web recommender systems” [17] se ha convertido en un clásico a la hora de distinguir los distintos tipos de SRs, en dicho artículo se proponen 6 enfoques de recomendación diferentes:

- Sistemas de recomendación basados en contenido: Este tipo de SR computan las recomendaciones basándose en las características de los ítems que el usuario ha valorado positivamente en el pasado.
- Sistemas de recomendación colaborativos (SRC): Estos sistemas usan las valoraciones de los usuarios para categorizarlos en grupos de acuerdo con la similitud de sus preferencias. Las recomendaciones se infieren teniendo en cuenta las valoraciones de los usuarios categorizados en un mismo grupo.
- Sistemas de recomendación demográficos: Estos categorizan a los usuarios en grupos demográficos basándose en atributos personales. Los usuarios, entonces, reciben recomendaciones de acuerdo con los grupos en los que se encuentran clasificados.
- Sistemas de recomendación basados en conocimiento: Estos sistemas infieren las recomendaciones basándose en las necesidades del usuario y en el conocimiento de las características de los ítems. Típicamente, computan las recomendaciones usando procesos de razonamiento basado en casos, por ejemplo: el usuario provee un ejemplo de ítem similar al que le interesa y el sistema, a partir de este conocimiento, infiere un perfil de ítem que permita encontrar los mejores ítems en el espacio de búsqueda.
- Sistemas de recomendación basados en comunidades: Este tipo de SR sugiere los ítems basándose en las preferencias de las “amistades” del usuario en cuestión. Esta técnica sigue el epigrama “Dime con quién andas y te diré quién eres” [18,19].

- Sistemas de recomendación híbridos. Este tipo de sistema se basa en la combinación de cualquiera de las técnicas mencionadas anteriormente. Un sistema híbrido que combina las técnicas A y B trata de aprovechar las ventajas de A para solucionar las desventajas de B.

De acuerdo con varias investigaciones [20,21], la aplicación de los sistemas de recomendación colaborativos es el enfoque más exitoso para realizar recomendaciones.

2.1 Sistemas de recomendación colaborativos

Como se había mencionado anteriormente los SRC computan las recomendaciones a partir de la clasificación de los usuarios en grupos con valoraciones similares. Luego, el problema de la recomendación puede ser definido como: la tarea de estimar la respuesta de un usuario para un ítem que es nuevo para él, basándose en la información histórica almacenada en el sistema. El tipo de respuesta usuario – ítem varía de una aplicación a otra y se puede categorizar de 3 maneras: escalar, binaria y unitaria. Las respuestas escalares son numéricas (ejemplo: valoraciones de 1 a 5 estrellas) u ordinales (ejemplo: Muy de acuerdo, De acuerdo, Neutral, En desacuerdo, Muy en desacuerdo) siempre representando los posibles niveles de apreciación de los usuarios respecto a los ítems. Las respuestas binarias solo tienen dos posibles valores donde se capturan dos niveles de apreciación opuestos (ejemplo: Me gusta, No me gusta). Finalmente, las respuestas unitarias capturan la interacción de un usuario con un ítem (ejemplo: compra, acceso online, etc.) sin brindar una información explícita de la apreciación del usuario por el ítem en cuestión.

Las maneras en las que las respuestas del usuario son capturadas pueden ser distintas también. Una manera es solicitar explícitamente una valoración al usuario pero otra manera podría ser el almacenamiento del historial de compras o acceso de un usuario a un grupo de ítems [22].

Se puede decir, entonces, que la idea central de los SRC es que la valoración de un usuario u para un nuevo ítem i será probablemente similar a la valoración de otro usuario v sobre i si u y v han valorado otros ítems de manera similar. El enfoque colaborativo resuelve alguna de las limitaciones del enfoque basado en contenido. Por ejemplo, los ítems para los cuales no se tiene información de sus características, o esta información resulta difícil de obtener, pueden ser recomendados a través de la retroalimentación provista por otros usuarios. Además, las recomendaciones obtenidas por un SRC se basan en la calidad de los ítems según el criterio de valoración de los pares del usuario activo en vez de basarse en el contenido, pues este puede ser un mal indicador de la calidad. Finalmente el enfoque colaborativo puede recomendar ítems con un contenido diametralmente distinto siempre y cuando otros usuarios hayan mostrado interés por estos ítems.

Los SRC automatizan el concepto de “rumor”, donde se confía en la opinión de personas con similitud de pensamiento u otra fuente confiable para evaluar la relevancia de un ítem (película, libro, etc.).

Existen 2 enfoques para realizar recomendaciones a partir de un SRC: “usuario - usuario” e “ítem - ítem”. El enfoque “usuario - usuario” se basa en que para obtener la predicción de la valoración de un usuario activo a cualquier ítem no valorado se haya la vecindad de ese usuario (usuarios que han efectuado valoraciones similares) para luego inferir la valoración que le daría al ítem el usuario activo. Por otro lado, el enfoque “ítem - ítem” se basa en que para obtener la valoración del usuario activo a cualquier ítem no valorado se haya la vecindad del ítem (ítems que han recibido valoraciones similares) y con esa información se predice la valoración que le daría el usuario activo. La selección de uno u otro enfoque depende básicamente en la estabilidad del grupo, por ejemplo, si los usuarios del sistema tienden a variar (se incorporan o se eliminan usuarios frecuentemente) se debe utilizar el enfoque “ítem - ítem”; si por el contrario la comunidad de usuarios del sistema tiene una mayor estabilidad sobre el conjunto de ítems ofertados por el sistema es recomendable utilizar el enfoque “usuario - usuario”.

Para cualquiera de los enfoques a utilizar dentro de los SRCs existen 2 métodos que permitirán obtener la predicción de la valoración del usuario activo sobre cualquier ítem no valorado por este. Estos métodos son el de regresión y el de clasificación. La opción de implementar el método de regresión o clasificación “usuario - usuario” depende, en gran medida, de la escala de valoración. Si el sistema utiliza una escala continua que va, por ejemplo, desde 1 a 10, es recomendable la utilización del método de regresión. Por el contrario si la escala de valoración contiene pocos valores discretos por ejemplo “Me gusta” y “No me gusta” o si los valores de la escala no pueden ser ordenados de una manera trivial, es recomendable la utilización el método de clasificación.

De manera general la predicción r_{ui} “ítem - ítem” del usuario u a ítem i por el método de regresión se calcula como el promedio de las valoraciones normalizadas y ponderadas dadas por los k usuarios más cercanos a u mediante la siguiente fórmula [23] :

$$r'_{ui} = h^{-1} \left(\frac{\sum_{j \in N_u(i)} w_{ij} h(r_{uj})}{\sum_{j \in N_u(i)} |w_{ij}|} \right) \quad (1)$$

donde w_{ij} es el coeficiente de similitud entre el ítem i (ítem del que se quiere predecir la valoración que daría el usuario activo) y el ítem j (uno de los vecinos de i) y $h(r_{uj})$ sería la normalización de la valoración dada por el usuario activo a j .

De manera análoga la predicción “ítem – ítem” por el método de clasificación se obtendría mediante la siguiente fórmula [23]:

$$r'_{ui} = h^{-1} \left(\operatorname{argmax}_{r \in S'} \sum_{v \in N_u(i)} \partial(h(r_{vj}) = r) w_{ij} \right) \quad (2)$$

donde la función $\partial(h(r_{vj}) = r)$ toma valor 1 si $r_{vj} = r$ y 0 en caso contrario.

2.2 Uso de los GIS con los SRs de productos turísticos

Un GIS es un sistema de software diseñado para permitir la entrada, el almacenamiento, la edición, la captura, el análisis y la salida de información geográficamente referenciada [24]. Los GIS son herramientas con un alto valor para los servicios turísticos porque los productos que se ofrecen en este sector normalmente pueden ser ubicados en un mapa.

El diseño de la interacción y la interfaz gráfica de un SR puede afectar profundamente el proceso de toma de decisiones del usuario. La utilización de un GIS como vía de representar las recomendaciones de productos turísticos obtenidas por un SR, brinda la ventaja de permitir al usuario la visualización geográfica de las entidades e ítems recomendados de una manera intuitiva, permitiéndoles, entre otras cosas, evaluar la distancia que existe entre ítems de los que puede convertirse en un consumidor. Por lo que al pensar en la implementación de un SR para productos turísticos, la utilización de un GIS para la presentación de los resultados se torna prácticamente obvia.

3 Técnica para la recomendación de productos turísticos en Cuba

Para obtener un sistema que recomiende ítems de manera efectiva se deben tener en cuenta cada una de las partes de este sistema, no solo el motor de recomendación. El objetivo de este trabajo es esbozar una técnica que permita la obtención de este sistema. Para lograr esto se parte de una primera regla que define una estructura general para el sistema y luego se definen un conjunto de reglas para cada uno de los componentes de esta estructura teniendo en cuenta las características contextuales del entorno turístico en Cuba. En la figura 1 se define la estructura general que tiene como componentes principales la interfaz gráfica de usuario, el selector de recomendación y el motor de recomendación.

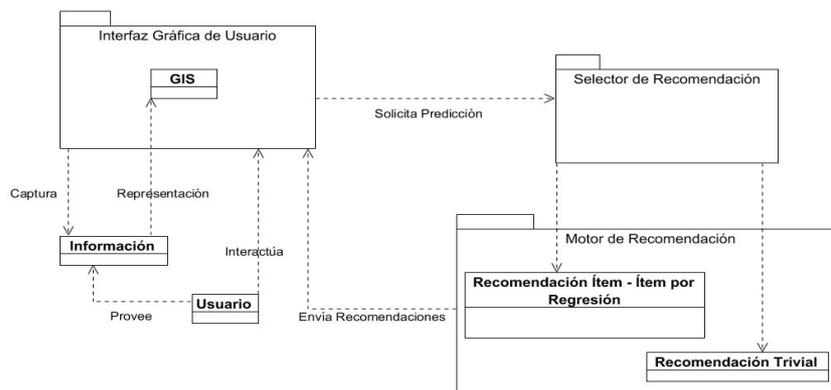


Fig. 1. Esquema del funcionamiento de la técnica para la recomendación de productos turísticos en Cuba.

3.1 Interfaz gráfica de usuario

La responsabilidad de este componente es la de garantizar la captura de la información que provee el usuario a partir de sus entradas de datos o de sus acciones sobre el sistema; además debe brindar una representación correcta de la información de manera que el usuario se sienta cómodo en su interacción con el sistema.

La información que debe ser representada es, básicamente, la referente a los productos turísticos que la organización que utilice el e-turismo desea ofertar a los usuarios. La manera en que esta información es presentada incide directamente en cómo los usuarios interactúan con el sistema así como el nivel de aceptación que estos le confieren. Para la representación de la información, luego, se propone utilizar un GIS debido a la característica inherente de los productos turísticos de poder ser ubicados geográficamente. El GIS debe mostrar un mapa de la región geográfica en cuestión donde se ubiquen los productos turísticos a ofertar y debe contar con funcionalidades básicas de zoom y paneo para permitir al usuario moverse libremente por el mapa.

Además, este componente debe representar la información referente a las recomendaciones que se le van a realizar al usuario, esta información constituye una entrada de este componente y se obtienen a partir del componente Motor de Recomendación. Se propone separar la visualización de las recomendaciones del mapa, es importante no efectuar recomendaciones sobre el mapa pues esto limitaría la variedad de productos con una buena predicción de valoración que pudieran ser recomendados cuando el usuario esté visualizando un fragmento de toda la región geográfica representada en el mapa.

La información que debe ser capturada, dando cumplimiento a la segunda responsabilidad de este componente, es aquella referente a la opinión del usuario sobre algún ítem así como aquella asociada a la descripción de la situación del usuario en el sistema. La valoración de los usuarios sobre los ítems que le son mostrados es la base del sistema de recomendación, estas valoraciones constituyen la información sobre la cual serán calculadas todas las predicciones que utilizará el sistema para decidir que ítems debe mostrar prominentemente a un usuario u otro.

Para el contexto de las organizaciones dedicadas al turismo en Cuba se propone la utilización de un método de captura de las valoraciones explícito y con una escala continua; esta manera debe implementarse de forma tal que no se interrumpa la navegación del usuario. Por la relación que guarda con la actividad turística y la manera en que son categorizados la mayoría de las instalaciones de alojamiento en Cuba se propone usar una escala de valoración de 1 a 5 esta escala podría ser presentada en forma de estrellas cuando el usuario esté visualizando la información detallada de algún ítem así éste podría valorar la calidad de un servicio o producto de 1 a 5 estrellas sin ver afectada, además, su interacción con el sitio como lo sería si se le mostrara un formulario adicional solo para que la valoración sea provista.

Por otro lado, la información situacional o contextual del usuario en un instante de tiempo determinado es la base que utilizará el componente Selector de Recomendación

para decidir qué tipo de predicción debe utilizar para obtener el listado de ítems que debe ser recomendado. En este componente debe implementarse la captura de la información relevante a ser usada por el selector de recomendación, esta información sería la referencia geográfica del centro del fragmento de la región geográfica que está visualizando el usuario, el nivel de zoom con que se está visualizando el mapa o si el usuario está visualizando o no los detalles de un producto turístico. La ubicación geográfica del centro de la porción de el mapa que está siendo visualizada y el nivel de zoom se utilizarán para completar, por cercanía, el conjunto de ítems que van a conformar la vecindad a partir de la cual se calculará la predicción de valoración para cada ítem, para luego obtener el listado de productos a recomendar; la información de si un usuario esta visualizando los detalles de un producto se utilizará para recomendar productos similares a este y, de esta forma, contribuir a mitigar el problema del arranque en frío. Es importante destacar que este tipo de información en particular no es persistente dado que describe una situación en un instante dado, esta información constituye una salida del componente Interfaz Gráfica de Usuario y entrada del componente Selector de Recomendación.

3.2 Selector de recomendación

La responsabilidad del selector de recomendación será la de analizar cada escenario en que se encuentre cada usuario en particular para decidir qué método de recomendación debe ser utilizado. Es en este componente donde se realiza la mayor parte de la tarea correspondiente a mitigar el problema del arranque en frío. El problema del arranque en frío viene dado cuando no se tienen suficientes valoraciones de un usuario o sobre un ítem de manera que no se puede realizar una predicción efectiva de valoración para uno y el cálculo de la vecindad para el otro.

De manera general, la información que sirve de entrada a este componente es la información contextual del usuario activo que se había mencionado anteriormente, referencia geográfica del centro de la región que está siendo visualizada, nivel del zoom y si está siendo visualizada la descripción detallada de un producto. A esta información contextual se le debe agregar la información de la cantidad de valoraciones que ha realizado el usuario activo, esta información aunque es situacional (puede variar de un instante de tiempo a otro) no la ofrece el componente Interfaz Gráfica de Usuario y debe ser obtenida a partir de los datos de las valoraciones almacenados por el sistema. La cantidad de valoraciones suficientes para ofrecer una predicción efectiva es un número que debe ser definido por aquel que se encuentre implementando la técnica pues depende de la situación particular (volumen de usuarios, ítems y valoraciones con que la organización debe trabajar). En cualquier caso, luego de definido el número, el trabajo con él es sencillo: si el usuario tiene las valoraciones necesarias se utiliza el método de recomendación ítem – ítem correspondiente, de lo contrario se debe utilizar un método de recomendación trivial.

Por otro lado, el dato de la referencia geográfica del centro de la región que está siendo visualizada debe ser utilizado, de ser necesario, para completar la vecindad de ítems

que permitirá calcular una predicción de valoración. Por ejemplo, se puede definir una distancia X y, de no encontrarse los k ítems valorados por el usuario activo que son necesarios para predecir una valoración a cualquier ítem no valorado, se completaría el número k de ítems utilizando aquellos que se encuentren dentro de un radio X a partir del punto que marca el centro de la región que se está visualizando y de los que se obtengan una mayor predicción de valoración para el usuario activo. Aplicar esta restricción puede incidir, en la calidad del funcionamiento del sistema de recomendación que se obtenga. Esto se debe a que a la relevancia que tiene el componente geográfico de los productos turísticos, por ejemplo: para un usuario que se encuentre hospedado o tenga reservación en un hotel, existe la probabilidad de que esté más interesado en los productos turísticos relativamente cercanos al dicho hotel. Por esta razón, se asume que si el usuario está visualizando el mapa con un nivel de zoom que se presume elevado es porque ese usuario tiene un interés especial en esa porción del mapa y luego tendría sentido usar este completamiento con ítems que se encuentren en una distancia razonablemente corta. Una vez más, el número que se defina para X y para el nivel de zoom depende de la situación en particular en la que se encuentre el equipo de desarrollo cuando esté implementando la técnica.

La información de si se está visualizando los detalles de un producto sirve para decidir que se va a recomendar en caso que no se tenga un número de valoraciones suficientes para efectuar una recomendación. Si se está visualizando la descripción detallada de un producto se puede asumir que este producto está recibiendo un nivel de atención significativo y en consecuencia se deben recomendar los productos que se encuentren en la vecindad inmediata (tengan un mayor coeficiente de similitud) de aquel que está siendo visualizado.

De manera general, en el cumplimiento de su responsabilidad, este componente, primero debe analizar si el usuario activo posee el número suficiente de valoraciones necesarias para realizar una recomendación. De no contar con las valoraciones suficientes se debe solicitar al Motor de Recomendación una recomendación trivial como lo sería el listado de ítems más populares (mejor valorados). En el caso de que el usuario activo no cuente con las valoraciones necesarias pero esté visualizando la información detallada de un ítem, se debe recomendar el listado con la vecindad inmediata de ese ítem. Por último, si el usuario cuenta con las valoraciones suficientes para obtener una recomendación, el componente debe analizar si el usuario está haciendo énfasis en una región específica del mapa para solicitar al Motor de Recomendación el listado de ítems a recomendar usando o no el completamiento de vecindad a partir de la distancia del centro de la región que está siendo visualizada.

3.3 Motor de recomendación

La responsabilidad del Motor de Recomendación es la de obtener el listado de ítems que serán mostrados al usuario prominentemente. Como su nombre lo indica, este componente es el responsable de hacer que la herramienta obtenida sea capaz de recomendar los ele-

mentos que son, presumiblemente, de interés para el usuario. Es en este componente donde se debe implementar el comportamiento derivado del razonamiento matemático usado para predecir la valoración que un usuario activo daría a un ítem no valorado, para luego mostrar prominentemente el listado de ítems que posean una predicción de valoración más elevada. Como se discutió anteriormente, se ha demostrado recientemente que el enfoque colaborativo es el más exitoso para efectuar recomendaciones al usuario. Principalmente porque ofrece varias y útiles ventajas y porque sus deficiencias pueden ser mitigadas con relativa facilidad. Usando este enfoque se persigue predecir las valoraciones, para los ítems de los cuales éstas no se conocen, utilizando las valoraciones que los usuarios de la comunidad ya han dado. Siguiendo esta línea de pensamiento, existen dos métodos para la implementación del enfoque colaborativo: el “Usuario - Usuario” y el “Ítem – Ítem”. En el contexto turístico cubano, la población de clientes (usuarios) es usualmente más voluminosa e inestable que la de los productos disponibles para el consumo (ítems) por esto se propone la implementación del método “Ítem – Ítem”.

Para cumplir con su responsabilidad, este componente debe tener disponible, como información de entrada, toda aquella asociada a las valoraciones otorgadas por los usuarios del sistema a los ítems disponibles. Además, debe tener disponible el componente geográfico de cada uno de los ítems, en caso que deba efectuar un completamiento de la vecindad usando la distancia de estos respecto a un punto (x;y). Este punto constituye la única información que deberá proveer el Selector de Recomendación al Motor de Recomendación, y lo hará solo en el caso que se desee obtener una lista de recomendación aplicando esta restricción en particular. En cualquier caso, la información de salida que debe proveer este componente es una lista con los elementos que deben ser mostrados prominentemente (recomendados) al usuario que será obtenida a partir del tipo de recomendación que haya solicitado, en particular, el Selector de Recomendación.

4 Pruebas

Para probar la técnica descrita se implementó un Sistema de Recomendación de Productos Turísticos aplicando las reglas definidas en esta. Luego se pobló una base de datos con 4000 usuarios y 1000 productos turísticos para, entonces, generar aleatoriamente valoraciones de esos 4000 usuarios sobre los 1000 ítems.

Para analizar la información que brinda como resultado la herramienta implementada se utilizaron indicadores de la calidad de las recomendaciones obtenidas como lo son la probabilidad de recomendaciones relevantes (Precisión) y la probabilidad de que los ítems relevantes sean recomendados (Recall).

4.1 Precisión y recall

La Precisión y el Recall se computan a partir de una tabla de 2 x 2 como se muestra en la tabla 1. El conjunto de ítems debe ser dividido en dos clases: ítems relevantes o no relevantes. En el caso de la técnica propuesta en el presente trabajo se considerarán relevantes para un usuario aquellos ítems que hayan obtenido una valoración de 4 o 5. Además se deben considerar los ítems que fueron recomendados (Seleccionados) al usuario, para el caso de la técnica se recomiendan todos aquellos que superen un umbral determinado. A su vez, en el ámbito de la implementación de la técnica, el valor umbral de predicción que debe superar un ítem para ser recomendado fue definido en 3.5 unidades.

Tabla 1. Datos de entrada para el cómputo de la Precisión y el Recall.

	Seleccionados	No Seleccionados	Total
Relevantes	Nrs	Nrn	Nr
Irrelevantes	Nis	Nin	Ni
Total	Ns	Nn	N

Luego, la precisión de las recomendaciones efectuadas a un usuario determinado, que representa la probabilidad de que un ítem recomendado sea relevante para el usuario, se obtiene como la división de la cantidad de ítems que fueron recomendados y son relevantes entre el total de ítems que fueron recomendados [24]:

$$P = \frac{N_{rs}}{N_s} \quad (3)$$

por otro lado el Recall representa la probabilidad de que un ítem relevante sea recomendado. Esta métrica se calcula como la división de la cantidad de ítems que fueron recomendados y son relevantes entre el total de ítems que son relevantes para el usuario [24] :

$$R = \frac{N_{rs}}{N_r} \quad (4)$$

aunque lo normal para un sistema de recomendación es predecir la valoración que le daría un usuario a cualquier ítem que este no haya valorado, a priori se puede observar que, para este tipo de prueba, se deben generar las predicciones de valoración para el conjunto de ítems valorados por cada usuario.

4.2 Resultados de las pruebas

Luego del cómputo de los valores de Precisión y el Recall para cada usuario, se obtuvo la media de estas métricas para el sistema usando solo el método de recomendación Ítem – Ítem por regresión, que fue implementado en el Motor de Recomendación. Luego se procedió a efectuar los mismos cálculos pero para cuando se utiliza la técnica como un todo, donde la diferencia en la recomendación marca la utilización del componente Selector de Recomendación perteneciente a la técnica propuesta. Las figuras 2 y 3 muestran una comparativa de estos resultados para Precisión y Recall del sistema respectivamente.

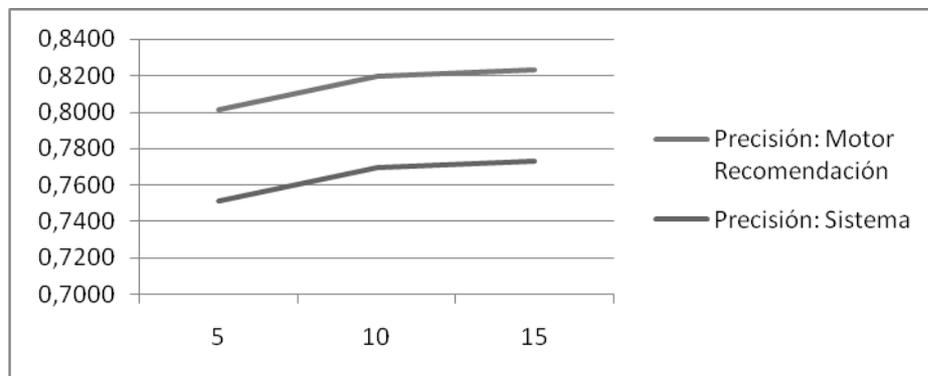


Fig. 2. Comparación de la precisión entre motor de recomendación y el sistema implementado.

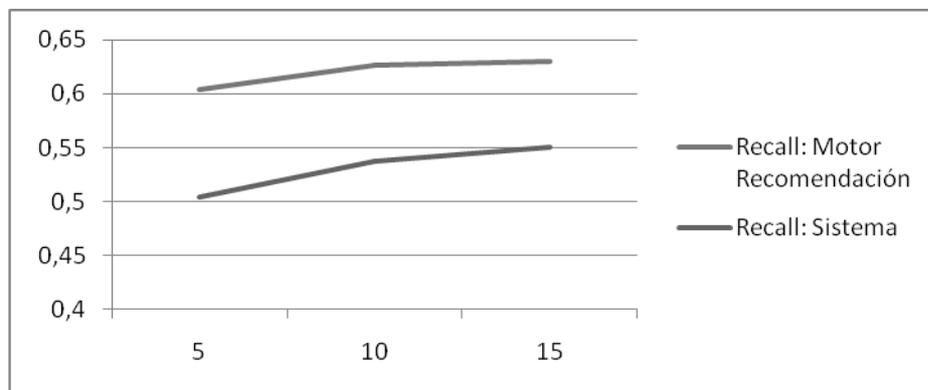


Fig. 3. Comparación del Recall entre motor de recomendación y el sistema implementado.

Como se puede apreciar, los resultados arrojados para la Precisión aumentan según aumenta el valor de k (5, 10, 15) para todos los casos. La probabilidad de que un ítem recomendado por el sistema sea relevante para el usuario, supera el 80 % para las 3 va-

riantes probadas en el caso del Motor de Recomendación en solitario y se mantiene sobre el 75 % cuando se utiliza el sistema como un todo. En el caso del Recall, la probabilidad de que un ítem relevante sea recomendado, también crece en relación directa con el valor de k y esta probabilidad se mantiene por encima del 60 % cuando se evalúa al Motor de Recomendación solamente y sobre el 50 % para la evaluación del sistema. Estos resultados se pueden catalogar de muy buenos para el sistema de recomendación obtenido pues permite afirmar que las recomendaciones obtenidas son probablemente útiles al usuario en cuestión. El decrecimiento de la Precisión y el Recall cuando se utiliza el sistema como un todo es un sacrificio aceptable teniendo en cuenta que, a priori, se pueden calcular un número significativamente mayor de predicciones de valoración en el espacio de predicción con lo que se logra una cobertura notablemente mayor.

5 Conclusiones

1. La importancia que, para Cuba, tiene la industria del turismo, hace de la personalización masiva de los contenidos que muestran las organizaciones de e-turismo cubanas una tarea necesaria.
2. La explotación de las características contextuales derivadas de la utilización de un GIS permite incidir positivamente en la cobertura del sistema sobre el espacio de predicción sin hacer grandes sacrificios en la precisión con que se calculan las predicciones de valoración.
3. La técnica propuesta define las características y el comportamiento que debe ser implementado para obtener recomendaciones de productos turísticos cubanos a partir de las preferencias y la información contextual de cada usuario.
4. Las pruebas realizadas demuestran que para el contexto turístico descrito se obtienen resultados positivos en la exactitud y el uso de las predicciones de valoración calculadas mientras se incide positivamente en la cobertura del sistema implementado, lo que permite afirmar que la aplicación de la técnica se logra obtener de manera eficaz el contenido que debe ser mostrado prominentemente a los usuarios de las organizaciones de e-turismo cubanas.

Referencias

1. Jannach, D., Zanker, J.M.M., Seidler, O.: Developing a conversational travel advisor with advisor suite. En: Information and Communication Technologies in Tourism, Proceedings, Frew, A.J., pp. 43–52, (2007)
2. Gao M. , Liu K., Wu Z.: Personalisation in web computing and informatics: Theories, techniques, applications, and future research, Information Systems Frontiers 12 607–629, (2010)

3. Adomavicius, G., Tuzhilin, A.: Towards the next generation of recommender systems: a survey of the state-of-the-art and possible extensions. *IEEE Transactions on Knowledge and Data Engineering* 17 (6), 734–749, (2005)
4. Resnick P., Varian H.: Recommender systems, *Association for Computing Machinery. Communications of the ACM*. 40 p. 56–58, (1997)
5. Charou, E., Kabassi, K., Martinis, A. and Stefouli, M.: Integrating Multimedia GIS Technologies in a Recommendation System for Geo-tourism. En: Tsihrantzis, G. A. and Jain, L. C. (eds.). *Multimedia Services in Intelligent Environment*, Springer-Verlag Berlin, pp. 63-74, (2010)
6. Werthner, H.: *Intelligent Systems in Travel and Tourism*. Sweet,P. (2001) *Designing Interactive Value Development*. PhD. Thesis, Lund University, Lund Business Press, Sweden.
7. Minghetti, V. and Buhalis, D.: Digital divide in tourism. *Journal of Travel Research* 49, 267–281, (2004)
- Chiang, C. and Jang, S.C. The effects of perceived price and Brand image on value and purchase intention: leisure travelers' attitudes toward online hotel booking. *Journal of Hospitality and Leisure Marketing* 15(3), 49–69, (2006)
8. Berka, T., Plobnig, M.: *Designing Recommender Systems for Tourism*. En: 11th International Conference on Information Technology in Travel and Tourism (2011)
9. Benckendorff, P.: An exploratory analysis of traveler preferences for airline website content. *Information Technology and Tourism* 8, 149–159, (2006)
10. Gretzel, U., Yuan, Y.L. and Fesenmaier, D.R.: Preparing for the new economy: advertising strategies and changes in destination marketing organizations. *Journal of Travel Research* 39, 146–156, (2000)
11. Ricci, F.: Travel recommender systems. *IEEE Intelligent Systems* 17, 53–55, (2002)
12. Kaplanidou, K. and Vogt, C.: A structural analysis of destination travel intentions as a function of web site features. *Journal of Travel Research* 45, 204–216, (2006)
13. Lee, J., Soutar, G. and Daly, T.: Tourists search for different types of information: a cross-national study. *Information Technology and Tourism* 9, 165–176. Kim, W.G., Ma, X. and Kim, D.J. (2006) Determinants of Chinese hotel customers' e-satisfaction and purchase intentions. *Tourism Management* 27, 890–900, (2007)
14. Mahmood, T., Ricci, F.: Improving recommender systems with adaptive conversational strategies. En: C. Cattuto, G. Ruffo, F. Menczer (eds.) *ACM, Hypertext*, pp. 73–82, (2009)
15. Law, R. and Jogaratnam, G.: A study if hotel information technology applications. *International Journal of Contemporary Hospitality Management* 17, 170–180, (2005)
16. Burke, R.: Hybrid web recommender systems. En: *The AdaptiveWeb*, Springer Berlin / Heidelberg, pp. 377–408, (2007)
17. Ben-Shimon, D., Tsikinovsky, A., Rokach, L., Meisels, A., Shani, G., Naamani, L.: Recommender system from personal social networks. En: K. Wegrzyn-Wolska, P.S. Szczepaniak (eds.) *AWIC, Advances in Soft Computing*, vol. 43, pp. 47–55. Springer.
18. Arazy, O., Kumar, N., Shapira, B.: Improving social recommender systems. *IT Professional* 11(4), 38–44, (2007)
19. Manouselis, N., Vuorikari, R., Assche, F.V.: Collaborative recommendation of e-learning resources: an experimental investigation, *Journal of Computer Assisted Learning* 26. p. 227–242, (2010)

20. Redpath, J., Glass, D.H., McClean, S., Chen, L.: Collaborative filtering: The aim of recommender systems and the significance of user ratings, en: ECIR , LNCS 5993, pp. 394–406, (2010)
21. Konstan, J.A., Miller, B.N., Maltz, D., Herlocker, J.L., Gordon, L.R., Riedl, J.: GroupLens: applying collaborative filtering to usenet news. *Communications of the ACM* 40(3), 77–87, (1997)
22. Ricci, F., Rokach, L., Shaphira, B., Kantor, P.B.: A Comprehensive Survey of Neighborhood-based Recommendation Methods. En: *Recommender Systems Handbook*, Springer New York / Dordrecht / Heidelberg/ London, pp. 117-130, (2011)
23. Demers, M.: *Fundamentals of Geographic Information Systems*. John Willey & Sons, (2008)
24. Herlocker, J. L., J. A. Konstan, et al.: "Evaluating collaborative filtering recommender systems." *ACM Transactions on Information Systems (TOIS)* 22(1): 5-53, (2004)