

EL MITO DE LAS OPINIONES MANIPULADAS EN TRIPADVISOR

Juan Pedro Mellinas Cánovas
losmellinas@yahoo.es
Universidad Politécnica de Cartagena

Juan Jesús Bernal García
juanjesus.bernal@upct.es
Universidad Politécnica de Cartagena

Soledad María Martínez María-Dolores
soledad.martinez@upct.es
Universidad Politécnica de Cartagena

RESUMEN

TripAdvisor es la mayor comunidad de viajeros online del mundo, pero la veracidad de su contenido es frecuentemente cuestionada, aduciendo un laxo sistema de control, que permite a cualquier individuo participar de forma anónima, dañando así el sistema. Este trabajo trata de disipar parcialmente las dudas de miles de hoteleros y turistas en todo el mundo, que se preguntan si es razonable confiar en lo publicado en esta web. Los resultados ponen de manifiesto que, para el caso de hoteles con una gran cantidad de opiniones, las puntuaciones de TripAdvisor resultan bastante fiables.

Palabras clave: TripAdvisor, opiniones, fraude, manipulación, hoteles

ABSTRACT

TripAdvisor is the world largest travel site, but the accuracy of its content is continually discredited, arguing a weak control system, which allows any individual to participate anonymously, damaging the system. This paper tries to clarify the doubts of many hoteliers and tourists around the world, wondering if it is reasonable to trust on this website. The results show that, in the case of hotels with a large number of reviews, TripAdvisor scores are quite reliable.

Keywords: TripAdvisor, reviews, fraud, manipulation, hotels

1. Introducción

Las opiniones que los consumidores vierten en internet sobre productos y servicios tienen una gran influencia en el resto de usuarios, pero cuando no se establecen estrictos sistemas de control, las empresas se pueden ver tentadas a incluir contenidos

positivos sobre sus marcas y/o negativos sobre la competencia. Los que emplean este tipo de técnicas tratan de defraudar al resto de miembros de la comunidad, adoptando diferentes identidades falsas que les permiten permanecer en el anonimato.

TripAdvisor, cuya relevancia en el sector turístico crece paralelamente a las controversias sobre la fiabilidad de sus contenidos, es probablemente el mejor ejemplo de esta nueva realidad. Según fuentes de la propia empresa (TripAdvisor, 2014), es “la mayor web de viajes del mundo” y “proporciona consejos fiables de viajeros reales”. Con casi 280 millones de usuarios únicos al mes, más de 170 millones de comentarios y opiniones de más de 4 millones de alojamientos, restaurantes y lugares de interés, no existen dudas sobre su relevancia dentro del mundo de los viajes. Sin embargo, no resulta tan clara la fiabilidad de la totalidad del contenido generado por sus usuarios, ya que cualquier internauta puede registrarse en esta web y aportar información sin tener que demostrar haber disfrutado del servicio sobre el que escribe.

TripAdvisor manifiesta contar con métodos de control que evitan la inclusión de contenido fraudulento, basadas en controles automatizados y manuales de los contenidos que entran en el portal (TripAdvisor, 2015). Pero los métodos de control que pueda imponer, simplemente consiguen dificultar la inclusión de opiniones falsas o limitar su número, aunque no evitarlas totalmente. En los últimos años han aparecido diversas noticias sobre casos de fraude en TripAdvisor en decenas de medios de comunicación de amplia difusión (New York Times, BBC, Daily Mail, etc.) sin que de esa información se pueda inferir un porcentaje exacto del nivel de opiniones manipuladas existentes. Estos hechos han contribuido a alimentar las teorías, suposiciones y conjeturas sobre la validez y fiabilidad del sistema dentro del sector hotelero.

En este contexto la revista Forbes (Forbes, 2012), publicó el artículo “TripAdvisor's \$35 Fair Value Could Suffer From Fake Review Concerns”, en el que hace referencia a la forma en la que la reputación de TripAdvisor se estaba viendo afectada por la publicación de noticias sobre la cantidad de críticas falsas presentes en su web. Se especulaba con que esta circunstancia podría afectar muy negativamente a la cotización de esta compañía, que para 2014 superaba los 10.000 millones de dólares de valoración en bolsa.

Algunas agencias de viaje online, también ofrecen la posibilidad a sus clientes, de opinar sobre su experiencia en los hoteles contratados. Este procedimiento dificulta la inclusión de opiniones fraudulentas, ya que sólo aquellos clientes que se han alojado y han pagado su estancia pueden aportar contenidos. Booking.com es el mayor exponente de este tipo webs, albergando la mayor base de datos de opiniones verificadas de viajeros en hoteles de España, junto a la ya mencionada TripAdvisor.

Tomando los datos de Booking como fiables y comparándolos con los obtenidos para una serie de hoteles en TripAdvisor, es posible realizar un acercamiento cuantitativo a la problemática de las opiniones fraudulentas. Los hoteles cuyas puntuaciones no sean similares en ambas webs, quedarían automáticamente señalados. Si esta circunstancia

se repitiera con una alta frecuencia podríamos cuestionar la fiabilidad del sistema en su conjunto.

2. La relevancia de las opiniones online y la propensión al fraude

Las opiniones de clientes en Internet son de gran ayuda para los consumidores (Chevalier y Mayzlin 2006), convirtiéndose en muchos casos en la principal fuente de información previa a la compra (Goldsmith y Horowitz, 2006.) Varios estudios demuestran que los individuos dan a este contenido generado por otros consumidores mayor credibilidad que la información que proporcionan críticos profesionales o las propias empresas (Chen y Xie, 2005).

El sector turístico y más concretamente el hotelero, son especialmente susceptibles de verse afectados por las opiniones vertidas por sus clientes (Buhalis y Law, 2008; Litvin *et al.*, 2012), siendo un perfecto ejemplo de eWom (electronic Word of mouth), que es la expresión inglesa del conocido como “boca a boca” o “boca a oreja” en español, adaptado a internet. Diversos estudios demuestran que las opiniones de viajeros en internet tienen gran influencia en las decisiones de los consumidores a la hora de elegir un establecimiento hotelero (Vermeulen y Seegers, 2009; Gretzel y Yoo, 2008). Un estudio dirigido por Dickinger y Mazanec (2008) demostró que las recomendaciones de amigos y las opiniones en internet son los dos factores que más influyen a los consumidores a la hora de elegir un hotel.

Los hoteles con mejores calificaciones, aumentarán sus probabilidades de ser elegidos por los consumidores y esto repercutirá en un aumento de los precios que pueden cobrar a sus clientes (Yacouel, 2012; Ogüt, 2012). Si el resto de factores permanecen constantes, unas buenas calificaciones del hotel, deberían suponer un aumento del precio que los consumidores están dispuestos a pagar por una habitación de hotel (Melnik y Alm, 2002).

La posibilidad de que parte de esas críticas puedan ser falsas, pone en grave riesgo la credibilidad de cualquier comunidad online y hace necesario el establecimiento de sistemas de control para evitarlo (O'Mahoney *et al.*, 2004). En 2012, la consultora Gartner anunciaba que para 2014 entre el 10 y el 15% de las opiniones en internet serían falsas (Sussin y Thompson, 2012). Yelp.com es una web norteamericana, con implantación internacional, que permite a sus usuarios realizar comentarios sin probar haber disfrutado del servicio sobre el que opinan, lo que puede propiciar la inclusión de opiniones falsas malintencionadas. Un estudio realizado en 2013 (Luca y Zervas, 2013) sobre opiniones de Restaurantes en esta web, arroja porcentajes del 16% de opiniones identificadas como fraudulentas.

El sector hotelero no es ajeno al mencionado fenómeno de las críticas falsas, siendo TripAdvisor el principal foco de la polémica (Yoo y Gretzel, 2009). El valor y relevancia de esta web no se basa solamente en los millones de usuarios y críticas que alberga, sino en la credibilidad de esa información, por lo que la sospecha de existencia de opiniones falsas pone en peligro esa credibilidad (Jeacle y Carter, 2011; Ayeh *et al.*, 2013). Cuando en 2008 la situación comenzaba a ser preocupante, O'Connor (2008)

concluye que, aunque existían indicios de manipulación de opiniones, su incidencia no era realmente significativa. Similares conclusiones obtiene otro estudio desarrollado usando una muestra de hoteles en Singapur (Chua y Banerjee, 2013).

Se han desarrollado modelos de detección de opiniones falsas, centrados en TripAdvisor, utilizando diversos criterios que señalan ciertas opiniones como “sospechosas” de ser falsas (Wu *et al.*, 2010a; Wu *et al.*, 2010b). También se ha analizado el fenómeno desde el punto de vista de la Psicología y la lingüística computacional (Ott *et al.*, 2011) tratando de buscar una metodología que pueda detectar el contenido fraudulento. Estos autores, de la prestigiosa Universidad de Cornell, desarrollaron un software capaz de detectar opiniones falsas sobre hoteles con una precisión del 90%, que actualmente puede ser utilizado de manera experimental a través de la web www.reviewskeptic.com.

Las posibles formas de fraude en TripAdvisor no se limitan a directivos de hoteles que escriben opiniones positivas sobre su establecimiento y/o negativas sobre los de la competencia. En ocasiones se requiere la colaboración de los empleados del hotel para evitar ser detectados o se contrata a empresas o personas especializadas en estas prácticas. Existen otras conductas, menos evidentes, pero que TripAdvisor también califica expresamente como fraudulentas (TripAdvisor, 2015) como son “*La solicitud de opiniones de forma selectiva únicamente de aquellos huéspedes que tuvieron una experiencia positiva*” o actuar “*Ofreciendo incentivos como descuentos, mejoras en la categoría de las habitaciones o cualquier otro trato especial a cambio de opiniones*”. Además, no sólo los competidores son susceptibles de incluir opiniones negativas deshonestas, empleados descontentos, clientes que buscan gratuidades en el hotel o clientes tan insatisfechos que escriben varias opiniones negativas con identidades diferentes, son otros agentes a tener en cuenta.

3. Comparación de puntuaciones en diferentes webs

Mayzlin *et al.* (2012) publicaron un artículo en el que comparaban las puntuaciones obtenidas en TripAdvisor y Expedia para un grupo de 2.931 hoteles de Estados Unidos. El estudio incide en la dificultad para introducir opiniones falsas en un sistema como el de Expedia, frente a la relativa facilidad para hacerlo en sistemas como el de TripAdvisor. En base a esto, utilizan el sistema de Expedia como referencia fiable asumiendo que, de no existir manipulación en TripAdvisor, las puntuaciones en ambas webs deben ser similares. En este caso, son capaces de detectar que los hoteles independientes o de pequeñas cadenas, así como los que tienen otros establecimientos cercanos, son más propensos a generar opiniones manipuladas. Estos hoteles tienen mayor número de puntuaciones extremadamente positivas en TripAdvisor que en Expedia. De igual forma los hoteles cercanos tienen mayor número de puntuaciones extremadamente negativas en TripAdvisor.

Expedia utiliza un sistema de opiniones verificadas, similar al de Booking, aunque con una escala diferente (Mellinas *et al.*, 2015). Para el mercado europeo, como se ha indicado anteriormente, resulta más recomendable el uso de Booking para este tipo de

estudios, debido al volumen de opiniones que registra. Por ello utilizamos esta web para realizar las pertinentes comparaciones, pero a diferencia del citado artículo, no nos centramos en la detección de puntuaciones extremas (positivas o negativas), si no que realizaremos un análisis de las puntuaciones globales de los hoteles.

HotelReporting es una empresa especializada en gestionar comentarios de los huéspedes de hoteles, que en Septiembre de 2012 publicó los resultados de un estudio (Hotelreporting, 2012) en el que comparaba las puntuaciones obtenidas por 230 hoteles de 4 estrellas, con más de 100 opiniones cada uno, en Booking.com y TripAdvisor. Los resultados indicaron que no existían diferencias significativas entre uno y otro sistema con solamente un 1,87% de diferencia entre ambos. Aunque las conclusiones obtenidas son ciertamente interesantes, este tipo de metodología no permite detectar que opiniones concretas son manipuladas, ni tan si quiera establecer un porcentaje hoteles con puntuaciones anómalas, simplemente pone de manifiesto las escasas diferencias entre ambos sistemas, cuando analizamos hoteles con un volumen de opiniones alto.

Entendemos que el uso de este tipo de metodologías, puede ayudar a aportar luz sobre un tema, alrededor del cual hay tantas dudas y sospechas. Nuestro enfoque, aunque basado en los mismos principios, se centra más en detectar aquellos valores en las puntuaciones que indican que se ha podido desarrollar una efectiva estrategia de fraude por parte del hotel o que un hotel ha sido víctima de una campaña de opiniones negativas falsas. Por lo tanto, el principal objetivo es determinar si existe un número significativo de hoteles que muestran una falta de concordancia entre sus puntuaciones en TripAdvisor y Booking.

4. Metodología y muestra

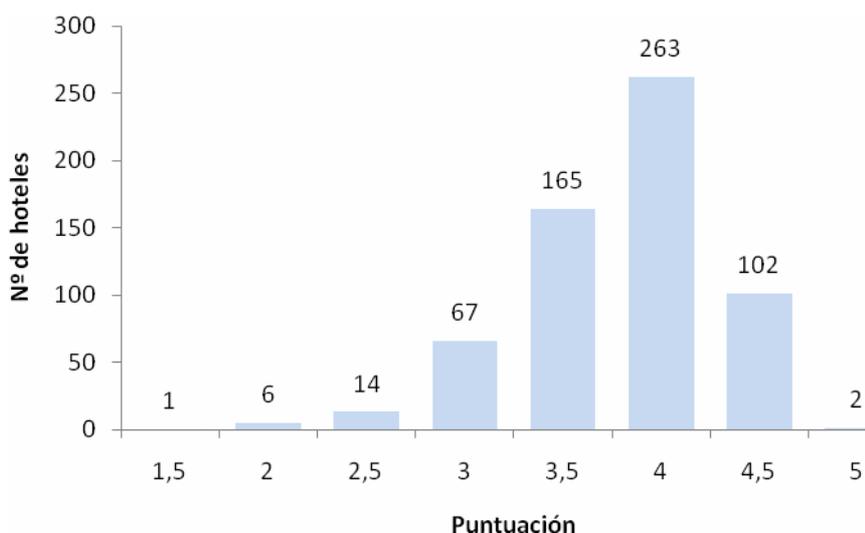
Ni TripAdvisor tiene un sistema infalible de detección del fraude, ni los estudios realizados al respecto son capaces de concluir que opiniones concretas son fiables, ni tan siquiera aportar una cifra concreta del porcentaje de opiniones manipuladas existentes. La metodología empleada en este estudio tampoco pretende obtener conclusiones definitivas al respecto, simplemente aportar una visión alternativa que propicie una mejor comprensión de la cuestión y facilite la identificación de anomalías.

Seleccionamos una muestra de hoteles, que nos permita acceder a un amplio número de establecimientos con un perfil lo más homogéneo posible. Por ello nos centramos en los hoteles dedicados al llamado turismo “de sol y playa” españoles, que probablemente alberguen el mayor volumen de plazas hoteleras a nivel nacional. Utilizamos una base de datos inicial constituida por los 1.440 hoteles de playa de las provincias mediterráneas e islas, utilizando el criterio de clasificación de Booking, que dentro de su “*Perfil del Hotel*”, tiene una categoría denominada “*Playa/Costa*”. También se selecciona la categoría “*Hotel*” dentro de la clasificación de Booking, evitando así hostales, apartamentos o albergues y filtramos para analizar únicamente hoteles de “*3,4 y 5 estrellas*”. Aplicando estos dos últimos filtros conseguimos cierto nivel de homogeneidad en la muestra.

Por otra parte, aunque TripAdvisor reconoce que, ocasionalmente, puede aparecer alguna opinión fraudulenta en su web, el gran número de opiniones registradas hace que esos posibles efectos se diluyan y la imagen global de los establecimientos responda a la realidad. Por ello descartamos hoteles con un número reducido de opiniones, en los que resulta fácil alterar su valoración, seleccionando aquellos que cuentan con un mínimo de 50 opiniones registradas en Booking, para seguidamente realizar una búsqueda manual de estos hoteles en TripAdvisor, seleccionando únicamente aquellos que también cuentan con un mínimo de 50 opiniones. De esta forma obtenemos una muestra de 660 hoteles con al menos 50 opiniones en ambas webs. Los datos fueron tomados durante el mes de Noviembre de 2011, una vez finalizada la temporada veraniega.

El sistema de valoración de TripAdvisor, aunque permite 9 puntuaciones distintas (de 1 a 5 puntos), concentra el 96% de las puntuaciones en sólo 4, para el caso de nuestra muestra (Figura 1). El hecho de que se opte por un redondeo limita la precisión del estudio, que ganaría en exactitud si se usara una escala más continua. Se podría haber optado por calcular una puntuación alternativa más precisa a partir de las frecuencias de todas las valoraciones en escala continua, que evitaría que puntuaciones tan diferentes como 3,3 y 3,7 acaben siendo redondeadas a 3,5, pero esta metodología tampoco estaría carente de limitaciones. TripAdvisor utiliza para el cálculo de la valoración global de un hotel un algoritmo secreto que da mayor valor a las opiniones más recientes, siendo esa puntuación la que realmente influye en la percepción de los clientes sobre el hotel, por lo que consideramos que utilizar este valor resulta más adecuado.

Figura 1
DISTRIBUCIÓN DE LA MUESTRA EN TRIPADVISOR



Fuente: Elaboración propia

Sin embargo, en el caso de Booking, se utiliza una escala con una puntuación que va de 2,5 a 10 (Mellinas *et al.*, 2015). En nuestra muestra encontramos puntuaciones que van de 5,6 a 9,4, aunque el 80% se concentren en el tramo 7-8,5. Adicionalmente el

hecho de que las opiniones de Booking.com sean incentivadas por el envío de un email al cliente y el mayor anonimato de que gozan aquellos que opinan en TripAdvisor (Gutierrez-Taño *et al.*, 2014), son condicionantes que pueden limitar la comparación de puntuaciones procedentes de ambos sistemas.

En todo caso, las divergencias entre las escalas de puntuación y la forma de captación de estas, deberían influir de igual manera en todos los hoteles. Es decir, si las puntuaciones con uno de los sistemas tienden a ser proporcionalmente más altas que con el otro, ese efecto debería repetirse en cualquier tipo de hotel. A mayores puntuaciones en TripAdvisor, corresponderán mayores puntuaciones en Booking, por lo que cualquier valor que no siguiera esta correspondencia, podría indicar la posibilidad de un fraude en el sistema. Nuestra hipótesis inicial sostiene que esos hoteles aparecerán claramente señalados, al comparar su puntuación en Booking, con la de otros hoteles con la misma puntuación en TripAdvisor.

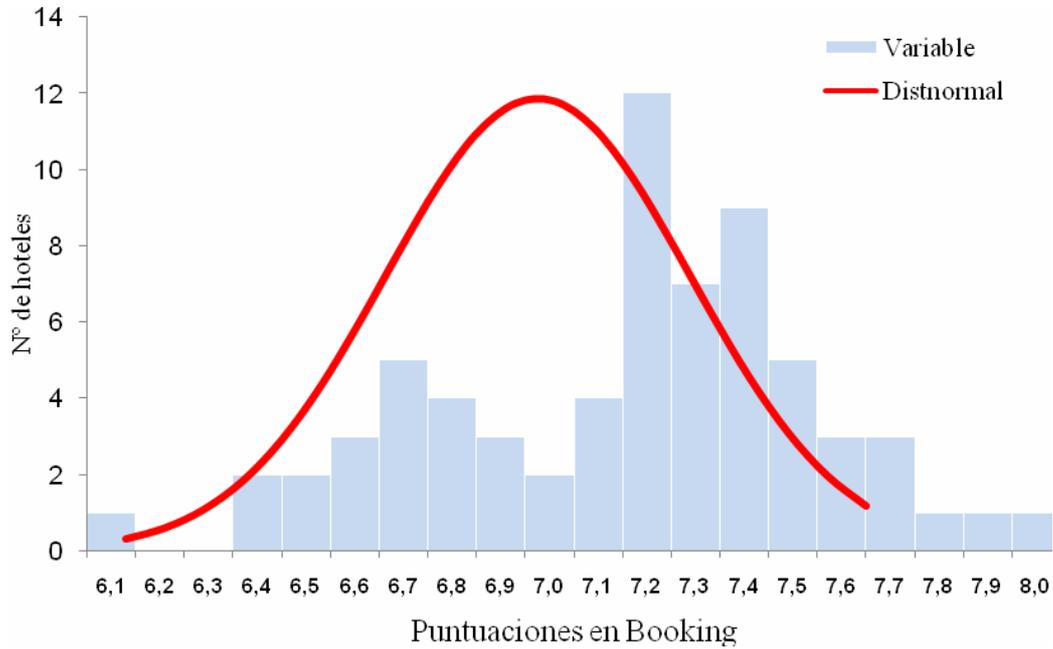
Analizamos las distribuciones de frecuencias de los hoteles de cuatro grupos de puntuaciones en TripAdvisor que representan el 96% de la muestra. Estos grupos cuentan con un número mínimo de hoteles, que pueden hacer cada una de las submuestras mínimamente significativas. Realizamos una representación gráfica de la distribución de frecuencias para cada uno de estos 4 grupos y calculamos una serie de variables estadísticas, que nos permitirán valorar los datos obtenidos.

La representación gráfica de la distribución de frecuencias sugiere que los datos son compatibles con una distribución normal para los citados 4 grupos, por lo que calculamos la distribución normal teórica resultante y realizamos el correspondiente test de normalidad para medir el nivel de ajuste real. Planteamos que la existencia de fraude o manipulación generalizada, junto a datos reales, no sería compatible con una distribución normal, ya que existirían un gran número de valores que no se ajustarían a este modelo. Los casos de puntuaciones manipuladas aparecerían en los extremos de las distribuciones de frecuencias, en un número tal, que no podríamos hablar de una distribución normal, tal y como se detalla en el apartado 6.

5. Resultados

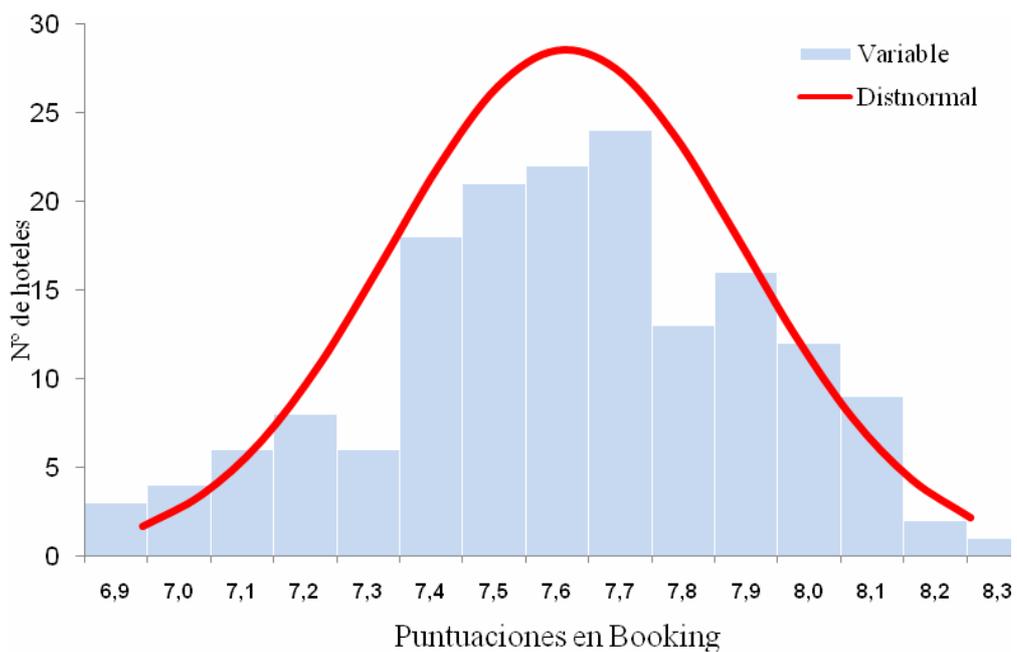
En las figuras 2, 3, 4 y 5 se muestran las distribuciones de frecuencias de los resultados obtenidos en cada caso, añadiendo la representación de la distribución normal teórica en cada uno de los casos.

Figura 2
HOTELES CON 3 ESTRELLAS EN TRIPADVISOR



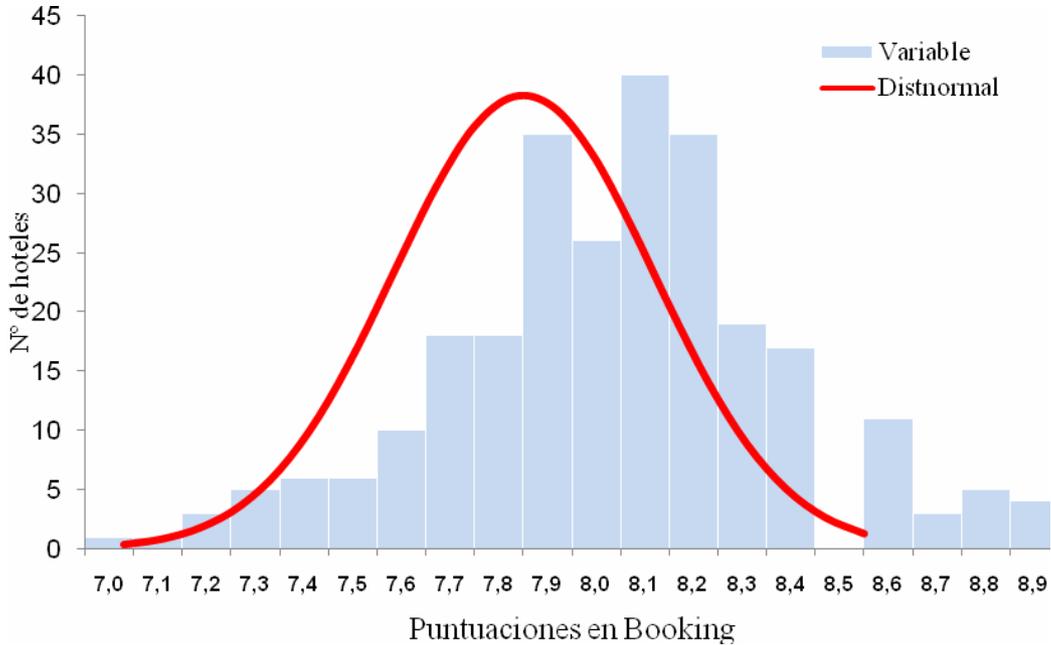
Fuente: Elaboración propia

Figura 3
HOTELES CON 3,5 ESTRELLAS EN TRIPADVISOR



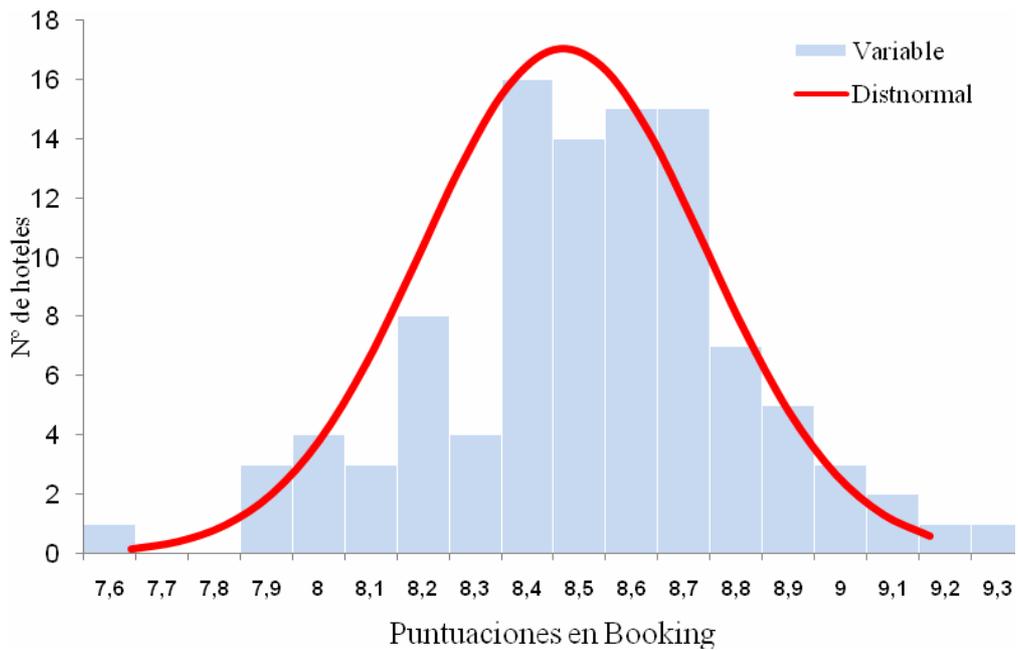
Fuente: Elaboración propia

Figura 4
HOTELES CON 4 ESTRELLAS EN TRIPADVISOR



Fuente: Elaboración propia

Figura 5
HOTELES CON 4,5 ESTRELLAS EN TRIPADVISOR



Fuente: Elaboración propia

Observamos que para la misma puntuación en Booking, puede existir una correspondencia con 2, 3 o incluso 4 categorías diferentes en TripAdvisor. Para una valoración de 8,4 en Booking (relativamente alta), tenemos prácticamente la misma cantidad de hoteles con 4 y con 4,5 en TripAdvisor. Un hotel con un 7,7 de puntuación en Booking, resulta bastante probable que tenga una valoración de 3,5 o 4 en TripAdvisor, aunque también encontramos casos, poco frecuentes, en los que tiene una valoración de 3. Para puntuaciones de 7,9 y 8, encontramos hoteles con las 4 diferentes puntuaciones de TripAdvisor analizadas, aunque la correspondencia más frecuente es la de 3,5 o 4.

Nos planteamos realizar el test de normalidad para contrastar la hipótesis nula (H_0) de normalidad de la población en los cuatro tipos de hoteles. Calculamos el estadístico de Prueba de Kolmogorov-Smirnov (KS), y lo comparamos con los valores críticos $D(n,\alpha)$ de este estadístico que están tabulados en la tabla correspondiente y comprobamos que se no se puede rechazar la hipótesis de normalidad en todos los casos, con un nivel de significación alto (20%) en tres de ellos. (Tabla 1).

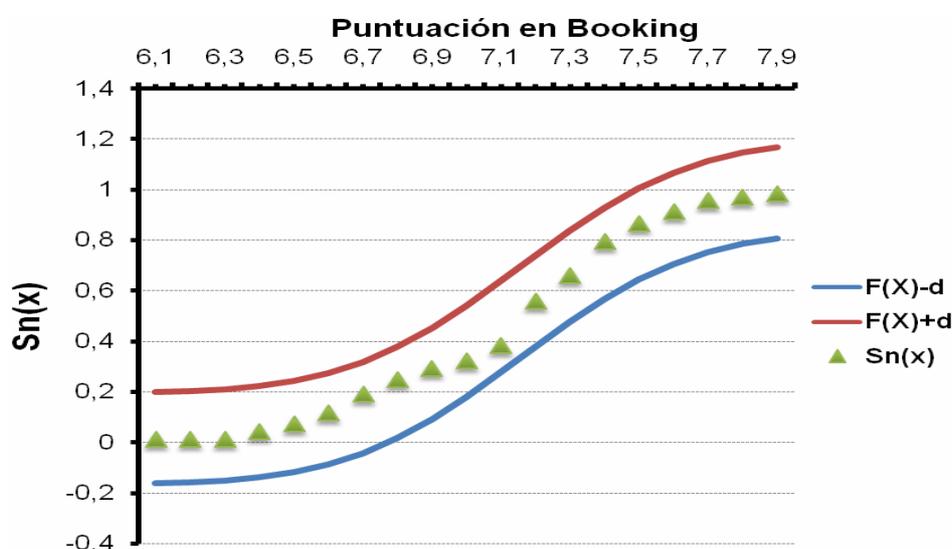
Tabla 1
DATOS ESTADÍSTICOS DE LAS DISTRIBUCIONES

	Media	Desviación	Asimetría	Dmáx (KS)	n	α	Dn, α	H_0
3,0 *	7,157	0,392	-0,384	0,06939011	68	0,2	0,12975656	SI
3,5 *	7,617	0,299	-0,231	0,06952327	165	0,2	0,08329939	SI
4,0 *	8,034	0,349	-0,253	0,0927112	263	0,02	0,09372722	SI
4,5 *	8,521	0,301	-0,220	0,08925278	102	0,2	0,10594579	SI

Fuente: Elaboración propia

A modo de ejemplo se adjunta un gráfico donde se aprecia de forma visual el resultado positivo del test de normalidad de Kolmogorov-Smirnov (Figura 6).

Figura 6
TEST DE KOLMOGOROV-SMIRNOV PARA HOTELES CON 3 ESTRELLAS



Fuente: Elaboración propia

La representación gráfica de los datos no indica una clara asimetría en ningún sentido, pero la aplicación del coeficiente de asimetría, indica una orientación de los datos hacia el lado izquierdo (valores negativos) en los 4 grupos. En todo caso, la observación de las propias gráficas, pone de manifiesto que se trata de una asimetría poco pronunciada.

6. Interpretación de los resultados

Estos resultados nos permiten elaborar una estrategia que facilite la determinación de los niveles de fraude existentes. Un hotel que genera opiniones positivas contraviniendo la política de TripAdvisor, obtendrá una puntuación media más alta de lo normal. Por ejemplo, un hotel con un 7,2 de puntuación en Booking, si actuando de forma correcta hubiera obtenido una puntuación de 3,5, entonces las acciones de fraude realizadas de forma exitosa deberían proporcionarle puntuaciones de 4 o 4,5.

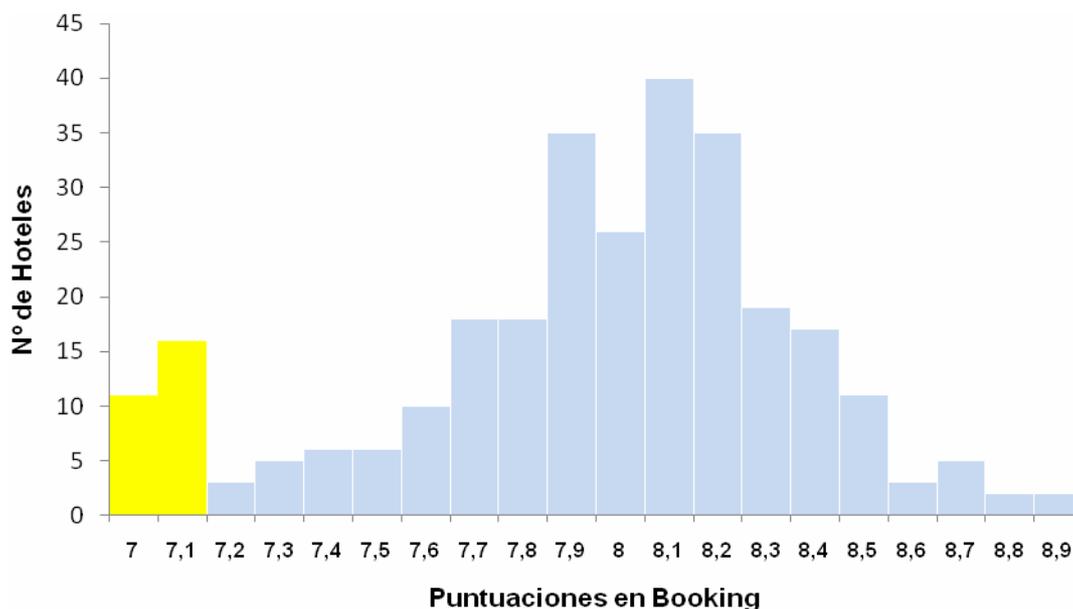
Por el contrario, si un hotel recibe opiniones negativas generadas de forma fraudulenta, obtendrá una puntuación sensiblemente más baja de lo normal en TripAdvisor, si esas opiniones negativas son suficientes y la estrategia de difamación tiene realmente éxito. Un hotel, con un 7,7 en Booking, al que le correspondería una puntuación de 4 si no existiesen irregularidades, podría ver bajar su valoración a 3,5 o 3.

Entendemos que si un hotel ha defraudado en TripAdvisor o ha sufrido un ataque, eso se verá reflejado en la relación de puntuaciones entre TripAdvisor y Booking. Así, los hoteles con éxito en su estrategia fraudulenta, aparecerán en el extremo izquierdo de la distribución de frecuencias para una puntuación concreta en TripAdvisor. Son hoteles que consiguen la misma puntuación que otros en TripAdvisor, pero partiendo de una puntuación mucho más baja que el resto en Booking. En este caso, la mayor parte de los hoteles con una puntuación similar en Booking, están una o dos categorías por debajo en TripAdvisor.

Si por el contrario, el hotel es víctima de una campaña de opiniones negativas fraudulentas, aparecerá en el extremo derecho de la distribución de frecuencias para su puntuación en TripAdvisor. Son hoteles que obtienen menor valoración en TripAdvisor de lo que se podría pensar por su puntuación en Booking.

Estos valores serían los llamados *outliers* en cada distribución, valores en los extremos de la distribución que pueden indicar la existencia de una irregularidad. Sin embargo, al observar las representaciones gráficas, parece que no existen como tales *outliers*, si no valores esperados en los extremos de una distribución normal. De existir un número medianamente significativo de hoteles con prácticas fraudulentas, deberían aparecer picos, como los señalados en amarillo en la figura 7, que correspondería con lo esperado en nuestra hipótesis inicial. Al observar las representaciones gráficas, parece que no existen tales valores atípicos, si no valores esperados en los extremos de una distribución normal

Figura 7
EJEMPLO DE SITUACIÓN CON HOTELES SOSPECHOSOS



Fuente: Elaboración propia

Consideramos lógico pensar que el fraude existente, corresponde principalmente a hoteles que utilizan estrategias irregulares para mejorar sus puntuaciones. Los casos en los de hoteles que sufren ataques de opiniones negativas por parte de la competencia o de otros agentes malintencionados deberían de ser menos frecuentes, sobre todo porque es más sencillo generar contenido positivo del propio hotel que contenido negativo de varios hoteles de la competencia. Resulta más beneficioso para un hotel generar 10 opiniones positivas sobre su establecimiento, que generar una opinión negativa en cada uno de los 10 competidores directos que tiene. Esta circunstancia, aunque de manera muy débil, parece ponerse de manifiesto en la asimetría detectada para los 4 grupos analizados.

Podríamos considerar que los valores situados en los extremos de cada gráfica no correspondieran a los esperados en una “distribución normal”, sino que fueran principalmente casos reales de fraude. Aún tomando aquellos casos que están en esos extremos y cuya puntuación en Booking es más propia de una categoría superior o inferior en TripAdvisor, su porcentaje sería mínimo, ya que no irían más allá del 5 o 10% del total de los hoteles.

Esta metodología no sirve, de por sí, para identificar los hoteles defraudadores o víctimas de ataques, aunque resulta perfectamente válida para descartar sospechas en ambos sentidos. Si las opiniones de un hotel presentan un patrón sospechoso de opiniones positivas, pero su valoración en TripAdvisor es coherente con la obtenida en Booking, habría que descartar la existencia de fraude o, al menos, la existencia de un número suficiente de opiniones manipuladas que afecten significativamente su calificación en TripAdvisor. Si un hotel argumenta ser víctima de una campaña de

difamación mientras tiene una puntuación en TripAdvisor, que coincide con la mayor parte de los hoteles de su misma puntuación en Booking.com, podríamos descartar las sospechas de un fraude masivo. El hotel ha podido sufrir de algunas opiniones negativas malintencionadas, pero quedaría demostrado que su escaso número, no ha afectado significativamente su puntuación en TripAdvisor. Por ejemplo, un hotel con un 7,5 en Booking, con una valoración de 3,5 en TripAdvisor, difícilmente puede ser acusado de generar opiniones fraudulentas de forma generalizada o de ser víctima de ataques de opiniones negativas.

Este mismo procedimiento resulta válido para identificar como “sospechosos” a los hoteles que se sitúan en los extremos de la gráfica. En nuestra muestra, encontramos dos ejemplos claros, que podría corresponder a este tipo de hoteles. Es el caso del hotel que con un 6,1 en Booking obtiene una puntuación de 3 en TripAdvisor y el del que con un 7,6 en Booking, obtiene un 4,5. En todo caso, habría que utilizar un criterio subjetivo, para determinar la amplitud del segmento de hoteles que se podrían etiquetar de esta forma.

Hay que señalar el hecho de que los hoteles de nuestra muestra tengan un mínimo de 50 opiniones registradas, estando la media por encima de las 100, puede contribuir decisivamente a que no se detecten divergencias entre las puntuaciones de Booking y TripAdvisor. A mayor número de opiniones naturales que tenga un hotel, mayor número de opiniones manipuladas se deben generar para que se produzca un efecto real y palpable en sus valoraciones. El trabajo que esto supone y la posibilidad creciente de ser detectado y castigado por TripAdvisor, puede contribuir a que los hoteles desistan de estas conductas.

También observamos que, cuando un hotel tiene una puntuación relativamente baja, el efecto de incluir opiniones muy positivas (5 puntos), tiene un mayor efecto que cuando el hotel ya tiene una puntuación alta. En la Figura 8 observamos un ejemplo de 4 hoteles con las 4 puntuaciones diferentes analizadas en nuestra muestra. Insertar 10 opiniones de 5 puntos en el caso del hotel de 3 puntos, le llevará probablemente a alcanzar una puntuación de 3,5 o incluso 4. Pero en el caso del hotel con 4 puntos, insertar el mismo número de opiniones manipuladas, es probable que no sea suficiente para alcanzar los 4,5 de valoración.

Figura 8
EJEMPLOS DE PUNTUACIONES EN TRIPADVISOR





Fuente: Capturas de pantalla de www.TripAdvisor.com

7. Conclusiones e implicaciones en el sector hotelero

Las decisiones metodológicas tomadas en el desarrollo del presente estudio, aunque entendemos que son adecuadas, implican la selección muestral de un número limitado de hoteles, de una tipología concreta y en un área geográfica determinada. Además, como se ha indicado, existen diferencias entre ambas escalas, que dificultan la comparación entre los sistemas de Booking y TripAdvisor.

En todo caso, consideramos que la metodología planteada resulta totalmente válida para su uso en la detección del fraude en TripAdvisor, especialmente en el señalamiento inicial de hoteles “sospechosos” de cometer irregularidades. Identificando los hoteles que se encuentran en los extremos de cada distribución de frecuencias, sería posible acotar grupos de establecimientos susceptibles de estar generando o recibiendo contenido manipulado. A estos hoteles, podrían aplicarse las metodologías anteriormente señaladas (Wu *et al.* 2010a; Wu *et al.* 2010b; Ott *et al.*, 2011) u otras alternativas, que permitirían determinar con mayor precisión la existencia de opiniones falsas.

La observación e interpretación de los datos obtenidos nos llevan a concluir que, aunque puedan existir de casos de opiniones manipuladas en TripAdvisor, los hoteles con una valoración sustancialmente alterada por esta razón parecen ser muy escasos. Un usuario de esta web, puede tener cierta seguridad de que un hotel con una alta valoración, en realidad no es un hotel mediocre que “haciendo trampas” ha conseguido una puntuación excelente. Así mismo los responsables de un hotel con una baja valoración, difícilmente pueden argumentar que esa situación es debida a ataques recibidos por agentes malintencionados. Esta conclusión se limita a hoteles con una cantidad mínima de opiniones registradas, lo que dificulta y hace poco atractivo establecer sistemas de manipulación. En los casos de hoteles con pocas opiniones, en los que la inclusión de un número reducido de opiniones extremadamente positivas, puede tener un efecto sustancial, la tentación de actuar de forma irregular puede resultar mucho mayor.

Cuando hablamos de una “valoración sustancialmente alterada” nos referimos a hoteles cuya puntuación es mucho mayor o menor en TripAdvisor de lo que sería si no actuara irregularmente. Los datos nos hacen descartar, casi por completo, los casos de hoteles que debieran tener 3 puntos de valoración y obtienen 4 o 4,5, así como los de hoteles que ven hundida su puntuación por ataques de opiniones manipuladas, con variaciones de un punto o más. Sin embargo, estos datos no descartan la existencia de hoteles que, actuando contrariamente a las reglas, aumentan ligeramente su puntuación, como sugiere la ligera asimetría detectada. Nos referimos a hoteles que deberían tener un 3,5 y obtienen un 4 o la de hoteles que por recibir opiniones negativas manipuladas bajan su puntuación en ese medio punto indicado.

Entendemos que para los hoteleros, este análisis puede servir para alcanzar un mayor conocimiento acerca de un tema sobre el que existen muchos “rumores y sospechas” y pocos datos fiables. La utilización de esta metodología, incluso de los datos mostrados, puede ayudar a disipar dudas sobre dos conjeturas recurrentes en el sector, cuando se habla de TripAdvisor:

- 1) Creer que el hotel propio no obtiene una mejor puntuación global porque una parte de las opiniones negativas recibidas provienen de agentes malintencionados.
- 2) Creer que los hoteles de la competencia obtienen mejor puntuación que el propio, no por satisfacer de mejor manera a los clientes, sino porque utilizan técnicas contrarias a las normas de TripAdvisor.

Para un hotelero, ignorar la información que proporciona TripAdvisor, argumentando la falta de autenticidad de gran parte de los contenidos que aparecen en esta web, no tiene sentido a la vista de estos datos. Aquellos que optan por dar la espalda a este tipo de información, tratan de negar una realidad, que no se ajusta adecuadamente a su percepción de la satisfacción de sus clientes.

Bibliografía

Ayeh, J. K., Au, N., y Law, R. (2013): “Do We Believe in TripAdvisor? Examining Credibility Perceptions and Online Travelers’ Attitude toward Using User-Generated Content”. *Journal of Travel Research*, Vol.52, No.4, pp. 437-452.

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*, Vol. 29, No.4, pp. 609-623.

Chen, Y., y Xie, J. (2008): “Online consumer review: Word-of-mouth as a new element of marketing communication mix”, *Management Science*, Vol. 54 No. 3, pp. 477-491.

Chevalier, J.A., y Mayzlin, D. (2003): “The effect of word of mouth on sales: Online book reviews”, *Journal of Marketing Research*, Vol. 43 No. 3, pp. 345-354

Chua, A. Y., y Banerjee, S. (2013): "Reliability of Reviews on the Internet: The Case of TripAdvisor". In Proceedings of the World Congress on Engineering and Computer Science (Vol. 1).

Dickinger, A., y Mazanec, J. (2008): "Consumers' Preferred Criteria for Hotel Online Booking", in Information and communication technologies in tourism 2008 proceedings of the international conference in Innsbruck, Austria, 2008, Springer-Verlag, Vienna, pp. 244-254.

Forbes (2012): "TripAdvisor's \$35 Fair Value Could Suffer From Fake Review Concerns", [Consulta: 30 de Julio de 2014]. Disponible en:
<http://www.forbes.com/sites/greatspeculations/2012/08/28/tripadvisors-35-fair-value-could-suffer-from-fake-review-concerns/>

Goldsmith, R. E., y Horowitz, D. (2006): "Measuring motivations for online opinion seeking", Journal of interactive advertising, Vol. 6, No.2, pp. 2-14.

Gutierrez-Taño, d., González-Wetherill, y.c. Y Parra-López, E. (2014): "La influencia de la autenticación de identidad en las valoraciones online de los hoteles". XVIII Congreso Nacional Asociación Española de Experto Científicos en Turismo (AECIT). Benidorm.

Gretzel, U. y Yoo, K.H. (2008): "Use and impact of online travel reviews", in Information and communication technologies in tourism 2008 proceedings of the international conference in Innsbruck, Austria, 2008, Springer-Verlag, Vienna, pp. 35-46.

Hotelreporting (2012): "¿Perjudican los comentarios falsos la puntuación final de los hoteles en TripAdvisor?" [Consulta: 30 de Julio de 2014]. Disponible en:
<http://blog.hotelreporting.com/actualidad-hotelera-y-mas.php/iperjudican-los-comentarios-falsos-de>

Jeacle, I. y Carter, C. (2011): "In TripAdvisor we trust: Rankings, calculative regimes and abstract systems", Accounting, Organizations and Society, Vol. 36 No. 4-5, pp. 293-309.

Litvin, S. W., Goldsmith, R. E., y Pan, B. (2008). "Electronic word-of-mouth in hospitality and tourism management". Tourism management, Vol. 29, No.3, pp. 458-468.

Luca, M., y Zervas, G. (2013): "Fake it till you make it: Reputation, competition, and Yelp review fraud". Harvard Business School NOM Unit Working Paper, (14-006).

Mayzlin, D., Dover, Y., y Chevalier, J. A. (2012): "Promotional reviews: An empirical investigation of online review manipulation" (No. w18340). National Bureau of Economic Research.

Melnik, M. I., y Alm, J. (2002): "Does a seller's ecommerce reputation matter? evidence from ebay auctions". The journal of industrial economics, Vol. 50, No.3, pp. 337-349.

Mellinas, J. P., María-Dolores, S. M. M., y García, J. J. B. (2015). "Booking. com: The unexpected scoring system". Tourism Management, Vol. 49, pp. 72-74.

O'Connor, P. (2008): "User-generated content and travel: A case study on TripAdvisor.com", in Information and communication technologies in tourism 2008 proceedings of the international conference in Innsbruck, Austria, 2008, Springer-Verlag, Vienna, pp. 47-58.

Öğüt, H., y Onur Taş, B. K. (2012). "The influence of internet customer reviews on the online sales and prices in hotel industry". The Service Industries Journal, Vol. 32, No.2), pp. 197-214.

O'Mahony, M., Hurley, N., Kushmerick, N. y Silvestre, G. (2004): "Collaborative recommendation: A robustness analysis", ACM Transactions on Internet Technology (TOIT), Vol. 4 No. 4, pp. 344-377.

Ott, M., Choi, Y., Cardie, C., y Hancock, J. T. (2011). "Finding deceptive opinion spam by any stretch of the imagination". In Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies, Vol. 1, pp. 309–319. Association for Computational Linguistics.

Sussin, j. y Thompson, e. (2012): "The Consequences of Fake Fans, 'Likes' and Reviews on Social Networks". [Consulta: 30 de julio de 2014]. Disponible en: <https://www.gartner.com/doc/2091515/consequences-fake-fans-likes-reviews>

TripAdvisor (2014): "Sobre TripAdvisor". [Consulta: 30 de julio de 2014]. Disponible en http://www.tripadvisor.es/PressCenter-c6-About_Us.html

TripAdvisor (2015): "Moderación de opiniones y detección del fraude". [Consulta: 17 de Septiembre de 2015]. Disponible en http://www.tripadvisor.es/vpages/review_mod_fraud_detect.html

Vermeulen, I.E. y Seegers, D. (2009): "Tried and tested: The impact of online hotel reviews on consumer consideration", Tourism Management, Vol. 30 No. 1, pp. 123-127.

Wu, G., Greene, D., y Cunningham, P. (2010a). "Merging multiple criteria to identify suspicious reviews". In Proceedings of the fourth ACM conference on Recommender systems, pp. 241–244.

Wu, G., Greene, D., Smyth, B., & Cunningham, P. (2010b). "Distortion as a validation criterion in the identification of suspicious reviews". In Proceedings of the First Workshop on Social Media Analytics, pp. 10–13. ACM.

Yacouel, N., y Fleischer, A. (2012): "The role of cybermediaries in reputation building and price premiums in the online hotel market". Journal of Travel Research, Vol.51, No.2, pp. 219-226.

Yoo, K.H., y Gretzel, U. (2009): "Comparison of deceptive and truthful travel reviews" in Information and communication technologies in tourism 2009 proceedings of the international conference in Amsterdam, The Netherlands, 2009, Springer-Verlag, Vienna, pp. 37-47.