

# Producing Recommendations by Analysing User Generated Content with the use of Fuzzy Logic and Multi-criteria analysis.

**Dr. Dimitris K. Kardaras**  
**Athens University of Economics and Business**

# Structure of presentation...

---

- Recommendation Systems (RS) in Tourism
- Introduction to User Generated Content (UGC)
- UGC in tourism
- Applying Fuzzy Logic and Multicriteria to analyzing UGC
- Future Research

# Recommendation Systems (RS) in Tourism

# Recommendation Systems-I

---

- A particular class of intelligent information assistants known as *recommendation systems* has been used in the travel domain, to provide travelers with relevant recommendations regarding their trips,
- E.g. which places to visit, which hotel to choose, ...etc.

# The tourism industry...I

---

- The World Travel and Tourism Council (WTTC) indicates that the contribution of travel and tourism to world GDP grew for the sixth consecutive year ...
- rising to a total of 9.8% of world GDP (US\$7.2 trillion).

# The tourism industry...II

---

- According to WTTC, the tourism sector employs 284 million people, which globally represents 1 in 11 jobs.

# Recommendation Systems-II

---

- Recommendation systems (RS) utilise techniques spreading from statistics, to AI and machine learning in order to capture user interests, build user and products/services profiles and suggest the most appropriate products or services to them.

# Recommendation Systems-III

---

- RS draw on several methods for developing user references models, with user-generated-content (UGC) to represent a source with rich customer information.



# Introduction to User Generated Content (UGC)

# UGC-I

---

- UGC refers to any type of content (e.g. blogs, website pages, images, social media posts, and testimonials) that is created and published by users (e.g. customers, fans).
- Users do not expect any type of payment as a return to publishing their reviews and comments.

# UGC-II

---

- Social media platforms allow users to exchange experience, feedbacks, opinions, complaints, etc., they **provide significant information** for capturing and understanding user interests.

# UGC-III

---

- Customers often express their experience by publishing their **reviews**.
- **Sentiment analysis** of user reviews provides the means for capturing and modelling users' preferences, emotions and attitudes, thus refining market segregation by grouping customers with similar needs and incentives and predicting customers' travel behavior more precisely.

# UGC in tourism

It is argued that there exist **four classes** of information contexts that need to be specified when attempting to understand user interests

---

# Four classes of information

---

- The *general information class* that refers to **personal characteristics** such as name, contact details, demographics of the user.
- The *event class* represents **user's activities**.

# Four classes of information

---

- The *preference class* refers to **user's interests**.
- The *social network class* explains **user's connections and interactions with other users**.



# Research show that ...

---

- while many systems produce and use user profiles, e.g. in web personalisation, recommender systems
- **there exists no definite procedure** for deriving user interests

# UGC in tourism-I

---

- Analysing UGC has been the focus of e-tourism research studies.
- Drawing on behavioural, socio-economic and demographic data analysis several researchers shed light into understanding people's travel behavior.

# UGC in tourism-II

---

- Surveys on travellers' preferences have shown that the travel selection process is complex depending among others, on personality and mood related factors, service quality issues, the Word-Of-Mouth (WOM) and the eWOM.

# UGC in tourism-III

---

- A variety of personality factors such as: personality traits, travel-related behaviours, travel experiences and feelings, lifestyle characteristics, and travel trends, determine the subjective assessment of travelling and tourism services.

# Applying Fuzzy Logic and Multicriteria to analyzing UGC

# Methodology steps I

---

1. **Select all documents** published by user(i).
2. **Identify the terms** that express user(i) preferences regarding tourism services that user(i) has consumed.

# Methodology steps II

---

3. **Calculate** the Term Frequency (TF) for each identified term.

# Methodology steps III

---

4. **Calculate the weight of each term** using the following formula:

$$W_{tk} = TF_{tk} * \log \frac{N_i}{d_{tk}}$$

Where  $W_{tk}$  represents the weight of term  $t_k$ ,  $TF$  is the term frequency,  $N_i$  is the total number of documents published by user (i) and  $d_{tk}$  represents the number of documents that contain term  $t_k$ .



# Methodology steps IV

---

5. Apply the **Fuzzy Analytic Hierarchy Process (FAHP)** multicriteria method, by taking into consideration the weight (importance) of each feature.

$$A = (a_{ij})_{n \times n}$$

# FAHP Steps

**Step 1:** Apply FAHP and construct the hierarchical model.

## Step 2: Assume the fuzzy pair-wise comparison judgment matrix

---

$$A^u = \begin{matrix} & t_1 & t_2 & \dots & t_k \\ t_1 & 1 & a^u_{(1,2)} & \dots & a^u_{(1,k)} \\ t_2 & \frac{1}{a^u_{(1,2)}} & 1 & \dots & a^u_{(2,k)} \\ \dots & \dots & \dots & 1 & \dots \\ t_k & \frac{1}{a^u_{(1,k)}} & \frac{1}{a^u_{(2,k)}} & \dots & 1 \end{matrix}$$

# The Pairwise Comparison Matrix of terms in users' reviews

---

...is formed by taking into consideration the weights of each term (feature).

So,  $a_{(1,2)}^u = \frac{w^1}{w^2}$ , where  $w^1$  and  $w^2$  indicate the weights of feature (1) and feature (2) respectively.

## Step 3: Formulas used in FAHP.

---

$$S_i = \sum_{j=1}^m M_{g_i}^j \otimes \left[ \sum_{i=1}^n \sum_{j=1}^m M_{g_i}^j \right]^{-1}$$

$$\sum_{i=1}^n \sum_{j=1}^m M_{g_i}^j = \left( \sum_{i=1}^n l_i, \sum_{i=1}^n m_i, \sum_{i=1}^n u_i \right)$$

$$\sum_{j=1}^m M_{g_i}^j = \left( \sum_{j=1}^m l_j, \sum_{j=1}^m m_j, \sum_{j=1}^m u_j \right)$$

## Step 4: Formulas used in FAHP.

$$\left[ \sum_{i=1}^n \sum_{j=1}^m M_{g_i}^j \right]^{-1} = \left( \frac{1}{\sum_{i=1}^n u_i}, \frac{1}{\sum_{i=1}^n m_i}, \frac{1}{\sum_{i=1}^n l_i} \right)$$

$$S_j = (l_j, m_j, u_j) \geq S_i = (l_i, m_i, u_i)$$

$$V(S_j \geq S_i) = \text{height}(S_i \cap S_j) = \mu_{S_j}(d) = \begin{cases} 1, & \text{if } m_j \geq m_i \\ 0, & \text{if } l_i \geq u_j \\ \frac{l_i - u_j}{(m_j - u_j) - (m_i - l_i)}, & \text{otherwise} \end{cases}$$

$$d'(A_i) = \min V(S_i \geq S_k), \quad \text{for } k = 1, 2, \dots, n \text{ and } k \neq i$$

$$W' = (d'(A_1), d'(A_2), \dots, d'(A_n))^T$$

# The FAHP calculates...

---

...the relative importance ( $P$ ) of each feature for each user. Thus, the preferences of user( $i$ ) are represented by vector ( $P$ ).

$p^{ui} = [p^k, p^l, \dots p^m]$ , where  $p^m$ , is the degree of preference of user( $i$ ) for feature ( $m$ ).



# Every user reviews refer to hotel services...

---

Each review expresses the sentiment of user (ui) for feature (y):  $S_y^{ui}$ .

Each review regarding a feature was published by user (ui) at a point in the time:  $T_y^{ui}$ .

# Both sentiment and time are fuzzified.

## Fuzzy Sentiment

Very Unpleasant	(0, 0.1, 0.25)
Unpleasant	(0.15, 0.3, 0.45)
Neutral	(0.35, 0.5, 0.65)
Satisfied	(0.55, 0.7, 0.85)
Very Satisfied	(0.75, 0.9, 1)

## Fuzzy Time

Very Old	(0, 0.1, 0.25)
Old	(0.15, 0.3, 0.45)
Neutral	(0.35, 0.5, 0.65)
Recent	(0.55, 0.7, 0.85)
Very Recent	(0.75, 0.9, 1)

Recommendation sets consist of top user's  
(ui) Evaluation of every feature (y) for each  
hotel.

Evaluations (E) are calculated by  
considering the preference (P), the  
sentiment (S) and the time (T)

---

$$E_y^{ui} = P_y^{ui} * S_y^{ui} * T_y^{ui}$$

# Evaluation of approach...

---

...The approach is tested with reviews (5,000 reviews) for hotels in London, UK, resulting an approximately 75% success rate of recommendations.

# Future Research

# Future research

---

- To specify the relative importance of Preference degree (P), Sentiment (S) and point in Time (T).
- To refine the formula used to estimate users' evaluation (E).
- Define different fuzzy sets for hotel categories.

**Thank you!!!**

---