

## EVALUATION CREDIBILITY OF OPINIONS ON A TRAVEL TOURISM WEBSITE

**Olimpia Ban, Versavia Ancusa**

*Economics Department, Faculty of Economic Sciences, University of Oradea, Oradea, Romania*

*Politehnica University of Timisoara, Timisoara, Romania*

[oban@uoradea.ro](mailto:oban@uoradea.ro)

[versavia.ancusa@upt.ro](mailto:versavia.ancusa@upt.ro)

**Abstract:** *When analyzing the online comments on some platforms, their perceived importance / accuracy and/or relevance relative to a location is quantified using two variables: the Gratitude – Appreciation Points (in case of amfostacolo.ro) and the score. While the score should be the “real” value, the Gratitude – Appreciation Points represents the social correction of that value. The research base was the travel comments website Am Fost Acolo/ I Was There (amfostacolo.ro), which is a Romanian site where you can gather and where you can post holiday impressions, you can watch pictures, you can see and compare deals and can make bookings.*

*Our hypothesis formulated and confirmed is: most of the opinions on a travel website are genuine. These results, in terms of generalization, have practical relevance for tourism managers.*

**Keywords:** satisfaction, accomodation, modeling online evaluation, credibility, amfostacolo.ro

**JEL classification:** Z33, C52, M31

### 1. Introduction

The content that expresses the consumer's experience and opinion about a product or service is known as "consumer-generated content or user-generated content" (CGC, UGC) This content creates word-of-mouth communication on the Web 2.0 platforms that is known in the specialty literature as the electronic word of mouth communication (eWOM). According to Yoo et al.(2011), eWOM is a new form of communication that serves the information needs of consumers by providing information: recent, detailed, non-commercial and it has the potential to reach beyond the consumer's social circles.

While traditional social relations require personal relationships to communicate, online social networks do not require this, enabling an impersonal communication, with less reliability disputes, being much more credible than advertisements (Chung and Buhalis, 2008).

Researchers in the travel and tourism industry have found that online reviews affect hotel rooms sales (Ye et. al, 2009; Vermeulen and Seegers 2009), and suggested that online reviews have higher levels of credibility than other sources of information (Dickinger, 2011).

eWOM is also a form of entertainment and interactivity as part of the products of services, absolutely necessary, as saying Pine and Gilmore (in Tonç and Tonç, 2016).

Research has also revealed that positive reviews improve attitudes toward hotels (Vermeulen and Seegers 2009) as well increase travelers' confidence and reduce their risk when booking accommodation (Gretzel, Yoo, and Purifoy, 2007 in Filieri and McLeay, 2013).

Ayeh, Au and Law (2013) have identified several directions discussing the UGC credibility:

- the credibility of those who write gives credibility to the messages posted. The readers are still facing the situation of reading the opinions of strangers (Park, Lee and Han, 2008; Litvin, Goldsmith and Pan, 2008);
- the different education and experience of those who write and who manage to write a more or less convincing message;
- the illicit intention of those who write with a manipulative role, under false identities. They do this to promote a hotel, to denigrate another, to get money etc. (Litvin, Goldsmith and Pan, 2008);
- entity managers who delete negative comments or write some positive comments;
- the credibility of the operators hosting these comments is questionable.

Travelers commonly use peripheral cues to assess the credibility of online information, such as the characteristics of the source and the message (Metzger et al., 2003).

However, the distrust of commercials makes many consumers to prefer the views of consumers just like them (Litvin, Goldsmith and Pan, 2008; Park, Lee and Han 2008) despite evidence that: the posts come from some individuals whom they do not see, do not know, do not know what school background they have, they do not know what expertise they have, or what degree of subjectivity or intentions they have.

DeAndreea et al.(2015): "People are often skeptical about the veracity of information appearing online because of the potential that exists for duplicity".

"The ability of website proprietors to control the dissemination of information generated by others was offered as an important factor that might affect how viewers judge the warranting value of information" (DeAndreea et al., 2015). 5% of product reviews came from an individual who never purchased an item from the retailer (Anderson & Simester, 2014 in DeAndreea et al., 2015).

From a data-analysis perspective, the multitude of available review informations represents an opportunity for automated analysis. However, every phase (Fisher et al., 2012) of the process has its challenges: the data acquiring phase (information sources selection), the data modeling phase (creating a suitable model and molding the data to it), the coding/debugging phase (dealing with large data sets creates computing difficulties) and the review phase (requiring a strong communication and understanding between specialists from different fields).

The data modeling phase can be extremely challenging, especially if the data is not designed to be so processed. While the model itself can have strong academic and practical confirmations, fitting the data onto it can rise another problem set. Data might be corrupt, malicious, with an unexpected structure or with particular curation rules.

## 2. Research methodology

Research objectives are to determine to what degree online comments and their assessment by other visitors, can be considered real and credible. Present issues that arise when modelling data that was not previously designed to be automatically analysed. Hypothesis formulated is:

H1: Most of the opinions on a travel web-site are genuine.

The research base was the travel comments website Am FostAcolo/ I Was There (amfostacolo.ro), which is a Romanian site where you can gather and where you can post holiday impressions, you can watch pictures, you can see and compare deals and can make bookings (Ban et al., 2015; Ban and Bădulescu, 2015). The website includes sites and accommodation facilities in over 60 destinations worldwide, with related evaluations and network moderators of the destination. The amfostacolo.ro website uses several indicators to evaluate the satisfaction, catching also the evaluation part beyond the 5 quality features used, that is:

- 5 features which are subject to scores from 1 to 10;
- the degree of satisfaction in percentages, given by the average of characteristics;
- the recommendation made by a reviewer;
- the appreciation points of the usefulness of the comment, awarded by the site visitors.

According to MiniGuide guide of AmFostAcolo.ro (<http://amfostacolo.ro/help9.php?id=38>), the organization system designed and implemented by AmFostAcolo is based on two fundamental concepts:

1. Sharing information (impressions, advice, recommendations) into two distinct categories:

- "accommodation" impressions - recommendations impressions about hotels, villas, guesthouses etc. and
- "travel" impressions - contain useful information, advice, recommendations about places worth (or not worth!) visiting (restaurant, towns, museums, belvedere places, beaches, mountains, national or local parks etc.).

2. The second important criterion is the geographical organization. All the "records" relating to a specific geographical area are "gathered into a" mini-library" bearing the sticker corresponding to the name of that area. The destinations include regions and the regions include sections.

The program calculates, based on the scores and recommendations of each review, two very important synthetic indicators for each hotel / villa / lodge:

- the average of scores awarded;
- the average degree of recommendation.

If the reviewer makes does not award any mark to a criterion, this criterion will automatically receive "-1" from the site administrators, and this criterion will be removed from average calculation.

Based on the above two indicators, the programme offers rankings of the most recommendable accommodation units, for each section, region or country.

The site uses the GAP system (Gratitude – Appreciation Points) ([http://amfostacolo.ro/pma\\_explic.php](http://amfostacolo.ro/pma_explic.php)).

Each information, text or photo uploaded on the site and in general every action useful to visitors brings a number of GAP. The management team evaluates the review and decides whether it is "accepted in the contest" (in this case it receives the GAP 1000 standard score, considered as "the vote of the site") or not. Each review admitted to the competition can get, in addition to the standard score, votes from the other users - APPRECIATIONS/ BONUSSES (with values of +450/ +900 GAP).

The impressions that are not admitted to the competition can only receive symbolic votes worth of +1 GAP ("dislike") or -1 GAP ("dislike"). Furthermore, points are awarded for the number of votes received, for photographs, photograph comments, replies to the posts etc.

The question behind this paper is How reliable are the GAP and score? What's the difference between them and how are they valued by the readers?

Data were gathered from the site on 04.30.2015 (posted at the time) and entered into an excel document, from where we selected the desired information as specified in the Table 1 (Ban et al., 2015).

**Table 1:** The characteristics of research population

<b>Characteristics</b>	<b>Absolute values</b>	<b>Relative values</b>
Total number of accommodation structures in Romania reviewed on the site	3755	100%
Number of reviewers for the structures in Romania by id/ by name:	9418/ 9417	100%
Total number of reviews for Romania	20883	100%
Number of reviews for locations in Romania	5683	27,21%
<b>Number of accommodation reviews for Romania</b>	15200	72,79%
Total number of reviews for hotels in Romania	5749	37,82%
Total number of reviews for pensions in Romania	5337	35,11%
Total number of reviews for villas and apartments in Romania	1329	8,75%
Total number of reviews for other accommodation structures in Romania	2785	18,32%
Age groups for reviewers for Romania		
<16 years old	32	0,34%
16-20	98	1,04%
20-30	2081	22,10%
30-40	4709	50,00%
40-50	1913	20,31%
50-60	441	4,68%
>60	88	0,93%
Type of travel of reviewers for Romania		
Single	182	1,93%

Childless couple	2973	31,57%
Families with children	5038	53,49%
Any	6	0,06%
Friends	1000	10,62%
Team-building	216	2,30%
Colleagues (there is this possibility in the file)	3	0,03%
Number of reviews according to the Gratitude-Appreciation Points awarded by the site administrators and by the other readers		
-700-0 (non-appreciation)	460	4,88%
1-2000	4926	52,30%
2000-3000	1257	13,35%
3001-4000	854	9,07%
4001-5000	786	8,35%
5001-10,000	864	9,17%
10,001-60,000	271	2,88%

(Ban et al., 2015)

### 3. Stages of research

In order to test the assumptions made, we started with the standardization of the corpus to be evaluated. Due to the large number of entries (15200 reviews) and to the stylistic variations of the inflows (8912 distinct users) there have been obvious difficulties in processing such as: the use of diacritics or not, the use of colloquial expressions, the use of abbreviations, vocalizations and so on. Therefore, we have created a mini-corpus of words that we have removed: linking words (and, with, the etc.), pronouns (I, me, that etc.), all the conjugations of auxiliary verbs to have and to be, regionalisms (is, mis, etc.), onomatopoeia / interjections (eeh, băi/yo, meh, etc), numerals (doi, doua, trei, etc /two, two, three, etc.), certain abbreviations (etc., and so on, pt/for) and misspellings (di, astia, abea, etc /di , guys, barely, etc). Some abbreviations have been replaced with literary forms of the words (eg "fff" with "very", "km" with "kilometers"). In total, the filtering corpus had over 480 distinct entries which amounted to 4,011,135 replacements in the corpus. In the end, this has allowed the normalisation of the evaluated corpus.

When analyzing the comments, their perceived importance/ accuracy and/or relevance relative to a location is quantified using two variables: the GAP and the score. While the score should be the "real" value, the GAP represents the social correction of that value. In other words, the social network corresponding to the "amfostacolo" site acts as a validator. In this case, it might be argued that even the number of votes per post can be relevant as a validation measurement. While the number of votes expresses how many people found the post helpful, this particular variable does not present the post quality, in the way GAP does.

As the social network is intrinsically spontaneously aggregated and , in this particular case, heterogenous (see Table 1) the validation represents not so much a "correctness" as a "perceived correctness" measure. Therefore, when automatically aggregating the data for a specific location, the disparaged variables should be compressed into an overall correct score (OCS).

In order to determine the underlining function of OCS, we'll analyze the variables that might impact it (Table 2).

**Table 2:** Value variation ranges for Votes, GAP and Score

	Votes	GAP	Score
Minimum value	1	69300	-1
Maximum value	79	-700	100
Average value	6.132829	3685	79.195 84
Average >0	6.132829	3822	79.985 24
Average <0	-	-48	-1
Count >0	15200	14660	15050
Count =0	0	74	149
Count <0	0	466	1

The first question we must answer is if the variables are independent. Therefore, the correlations between these three variables are presented in Table 3.

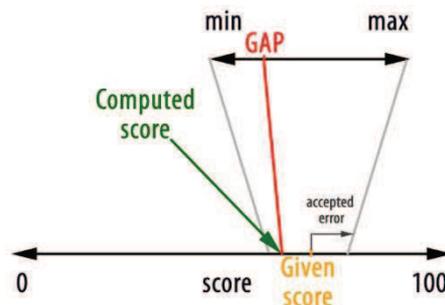
**Table 3:** Correlation between Votes, GAP and Score

	Votes	GAP	Score
Votes	1		
GAP	0.95707	1	
Score	0.152592	0.137195	1

The data presents a very strong correlation between GAP and Votes, therefore the OCS function should take into account only one of the “social-influenced” variables.

Referring to the working hypothesis, this indicates that the influence of network moderators is minimal, as the votes given by the social network are closely correlated with the GAP, which might be influenced by the network moderators through their competition model and point awarding. Therefore we consider hypothesis H1 as true.

As a future development guideline for the site we suggest the useage of OCS in which GAP acts as a “validator” for score.



#### **Figure 1: Scaling score according to GAP**

One question that we must answer before proceeding is the size of the individual error. For this particular case, as there haven't been reported any potential malicious postings, we considered a  $\pm 10\%$  an adequate accepted error range.

#### **4. Conclusion**

The use of online platforms for posting comments by current and potential tourism consumers is a widespread practice and with direct and measurable implications for tour operators. The problem that has arisen is the credibility of these comments and their isolation against false or malevolent comments.

In this paper, we suggest ways to investigate the credibility of the comments posted in terms of source (they are issued by individual commentators!).

The conclusion of the current study, carried out on the Amfostacolo.ro site have showed that most of the opinions on a travel web-site are genuine, and in this case, analysis may be reproduced for other platforms comments.

The proposed analysis has practical implications for communication platforms managers and managers of accommodation assessed on these platforms, because it can verify the veracity of comments posted here .

Also, this paper proposes an improved way of computing the real score, taking into account the social bias.

The limitations of the present research are related only to the testing of the analysis ways on only one travel site, which which prefigures also possible future research on other similar platforms.

#### **Acknowledgments**

We would like to express our gratitude for the support in this research to Mr. Cornel Bociort, amfostacolo.ro the site developer.

#### **References**

- Ayeh, J.K., Au, N. and Law, R. (2013) "Do We Believe in TripAdvisor? Examining Credibility Perceptions and Online Travelers' Attitude toward Using User-Generated Content", *Journal of Travel Research*, 52(4) 437–452.
- Ban, O., Bădulescu, A. (2015) "Evaluarea online a calității serviciilor de cazare și satisfacția globală // The online evaluation of accommodation services quality and the overall satisfaction", *Revista Română de Marketing*, nr.4\2015, pp. 60-70.
- Ban, O.I., Ancusa, V., Bogdan, V., Tara, I.Gh. (2015) "Empirical Social Research to Identify Clusters of Characteristics that Underlie the Online Evaluation of Accommodation Services", *Revista de Cercetare si Interventie Sociala*, 50, 293-308.
- Chung, J. Y., D. Buhalis. (2008) "Web 2.0: A Study of Online Travel Community." In *Information and Communication Technologies in Tourism 2008*, edited by P. O'Connor, W. Höpken, and U. Gretzel. New York: Springer Wien, pp. 70-81.

- DeAndrea, D.C., Brandon Van Der Heide, Vendemia, M.A. and Mao H. Vang, (2015) "How People Evaluate Online Reviews", *Communication Research*, 1–18, DOI: 10.1177/0093650215573862.
- Dickinger, A. (2011) "The Trustworthiness of Online Channels for Experience- and Goal-Directed Search Tasks", *Journal of Travel Research*, 50 (4): 378-91.
- Filieri, R., McLeay, F. (2013) "E-WOM and Accommodation: An Analysis of the Factors That Influence Travelers' Adoption of Information from Online Reviews", *Journal of Travel Research*, 53(1) 44–57.
- Fisher, D., DeLine, R., Czerwinski, M. and Drucke, S. (2012) "Interactions with Big Data Analytics", *Interactions*, pp. 50-59, May - June 2012, DOI: 10.1145/2168931.2168943.
- Litvin, S., Goldsmith, R., Pan, B. (2008) "Electronic word-of-mouth in hospitality and tourism management", *Tourism Management*, 29, 458–468.
- Melián-González, S. Bulchand-Giduma1, J. & López-Valcárcel, B.G. (2013) "Online Customer Reviews of Hotels: As Participation Increases, Better Evaluation Is Obtained", *Cornell Hospitality Quarterly*, 54(3) 274–283.
- Metzger, MJ, Flanagin AJ and Zwarun L (2003) "College student web use, perceptions of information credibility, and verification behavior", *Computers & Education*, 41: 271–290.
- Park, D., J. Lee, and I. Han. (2007) "The Effect of On-line Consumer Reviews on Consumer Purchase Intention: The Moderating Role of Involvement", *International Journal of Electronic Commerce*, 11 (4): 125-48.
- Tonț, D.M., Tonț, M.D. (2016) "An Overview of Innovation Sources in SMES", *Oradea Journal of Business and Economics*, Volume I Issue 1 Published on 31 March 2016, pp. 58-67.
- Vermeulen, I. E., Seegers, D. (2009) "Tried and tested: The impact of online hotel reviews on consumer consideration", *Tourism Management*, 30, 123-127.
- Ye, Q., R. Law, and B. Gu. (2009) "The impact of online user reviews on hotel room sales", *International Journal of Hospitality Management* 28:180-82.
- Yoo, K., Gretzel, U., Zach, F. (2011) "Travel Opinion Leaders and Seekers", *Information and Communication Technologies in Tourism: Proceedings of the International Conference*, New York Springer, 525-535.
- Yoo, Kyung-Hyan, Yoonjung Lee, Ulrike Gretzel, and Daniel R. Fesenmaier. (2009) "Trust in Travel-Related Consumer Generated Media." In *Information and Communication Technologies in Tourism 2009*, edited by Wolfram Höpken, Ulrike Gretzel, and Rob Law. Wien: Springer, pp. 49-59.