ANALYSIS OF SOCIAL VALIDATION BEHAVIOUR ON TRAVEL ONLINE REVIEWS WEBSITES

Olimpia Ban, Versavia Ancusa
Economics Department, Faculty of Economic Sciences, University of Oradea, Oradea, Romania
Politehnica University of Timisoara, Timisoara,Romania
oban@uoradea.ro
versavia.ancusa@upt.ro

Abstract: *Psychological word-of-mouth marketing substrate as well as online user-generated content presents the need for validation of opinions / actions of individuals as consumers .*
Satisfaction with a comment posted can be done by likes , a written positive answer or other forms , such as system the Gratitude – Appreciation Points, which 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 are: on a large enough scale, spontaneous social aggregation selects the relevant informations for the group and social validation has a predictable behaviour in time.These results, in terms of generalization, have practical relevance for tourism managers.

Keywords: e-world-of-mouth, accomodation, modeling online evaluation, website, amfostacolo.ro

JEL classification: Z33, C52, M31

1. Introduction

Word-of-mouth communication on the Web 2.0 platforms are known in the specialty literature as the electronic word-of-mouth communication (eWOM). eWOM is a relative 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 (Yoo et al., 2011).
In all the phases of the holidays , the most trusted sources for consumers proved to be private sources . Personal sources are made up of friends, relatives and acquaintances who have used similar products. In the online environment, knowledge has expanded that limited sphere to knowledge from unkown people who utilize / are interested in the same products. On the other hand, those who do post an oppinion have a variety of reasons : to validate their own consumption , to prove a level of expertise as a form of gratitude / revenge against the bidder, to reduce cognitive dissonance post consumer, etc.
While traditional social relations require personal relationships to communicate, online social networks do not require this, enabling an impersonal communication

yet reliable, 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).
One important area to consider is the perceived usefulness for comments, reflected in their positive / negative or non-assessment by other review readers . Validation messages is in relation to their positive or negative character (Hauser et al., 1993).
"The concept of negativity bias suggests that people give more importance to negative information than positive information and hence consumers should give more attention and importance to negative WOM than pozitive WOM" (Ghosh, Varshney, Venugopal, 2014).

In modeling a system to process such data there are several phases (Fisher et al., 2012), each with its challenges: the data acquiring phase (information sources selection), the data modeling phase (creating a suitable model and molding the data to it), the codding/debuging phase (dealing with large data sets creates computingdifficulties) 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 curration rules.

2. Research methodology

The objective of this research is to verify the degree of social validation of posts , taking into account the proposed evaluation system Amfostacolo.ro site .
Present issues that arise when modelling data that was not previously designed to be automatically analysed.
Hypothesis formulated are:

H1: Social validation has a predictable behaviour in time.

H2: On a large enough scale, spontaneous social aggregation selects the relevant information for the group.

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).

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/ BONUSES [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.

Data were gathered from the site on 04.30.2015 (posted at the time) andentered 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

Characteristics	Absolute values	Relative values
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%
Number of accommodation reviews for Romania	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 variation 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 (îs, 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 on the body to be processed. In the end, this has allowed the standardization of the corpus to be evaluated.

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 or social validation of the post.

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.

In order to compute OCS, as stated before, we considered the GAP as a "validator" for *score*.

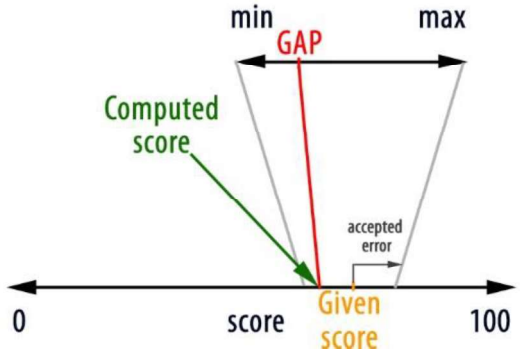

Figure 1: Scaling *score* according to *GAP*

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 ±10% an adequate accepted error range.

A problem with this type of scaling is that the GAP is presumed to be linear. Reffering strictly to the GAP variable, the minimum, maximum and average value,

suggest the fact that its scale is not linear, however, the hystogram for that value shows it's not pure exponential either (Figure 2)

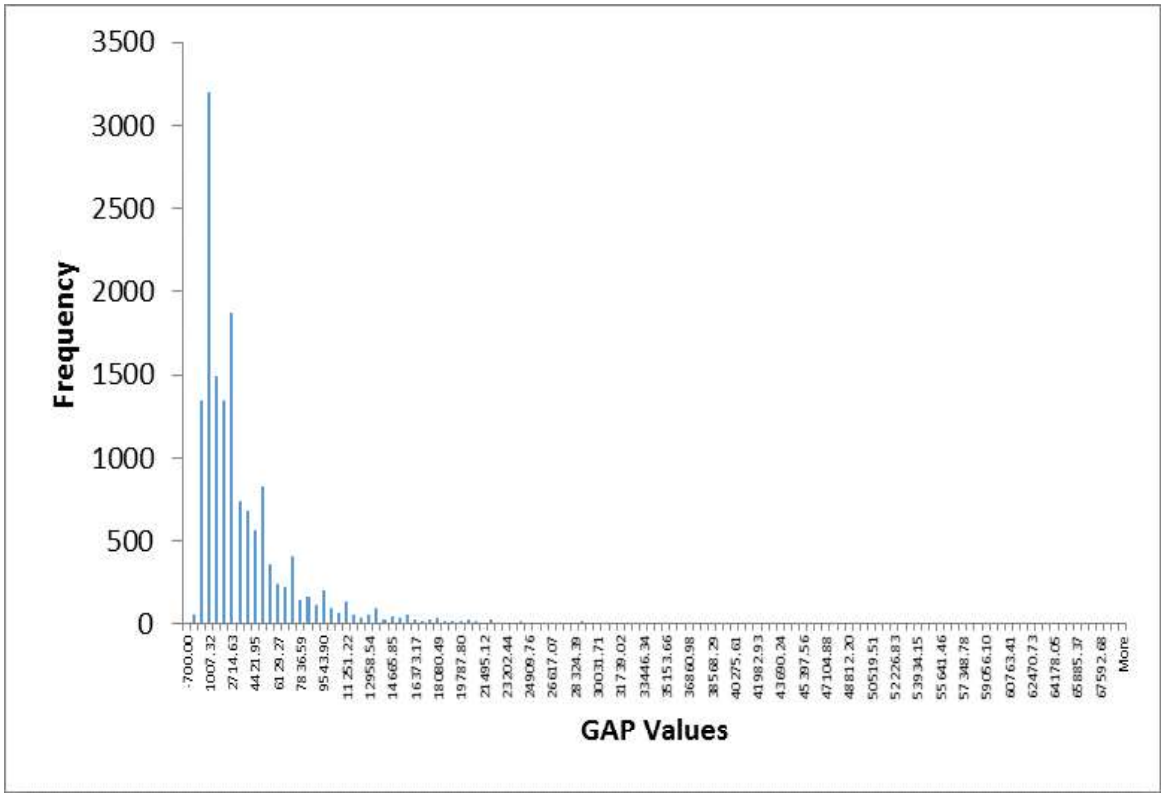

Figure 2: GAP hystogram

Regarding the matter strictly from a value-centric view-point, the positive and negative values of GAP have different variation forms: the negative is linear (Figure 3a) and the positive can be approximated with a power-law distribution (Figure 3b)

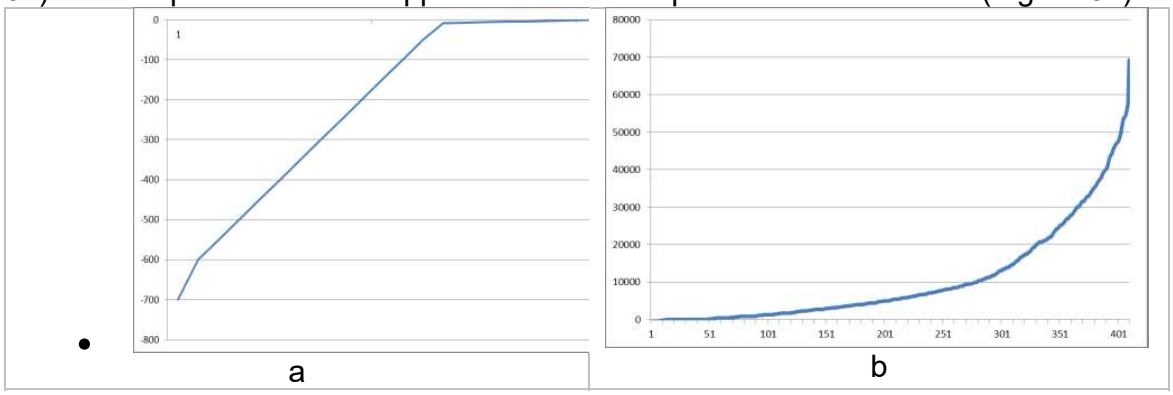

Figure 3: GAP range

If we want to use the GAP in order to scale the *score* that we must scale taking the GAP value variation form into account. In order to do that, the following steps were taken:

Step 1: Clean data

Due to the competiton model implemented by the site the ±1 symbolic scores lead to odd values, like 51 or -101 in the GAP column. In order to create a uniform value

set, these symbolic scores needed to be excluded, therefore we applied the following transformation. This gives us only 327 unique values for GAP, instead of 410 values that include the contaminatedones – it's a 20.24% reduction in the value set. However, from the histogram point-of-view the difference is minimal (see Figure 5, value series 1 and 2). The loss in GAP is felt only by the weakly appreciative GAP, the strongly appreciative GAP are not impacted. Through this we ensure that the most validated oppinions are put first, reflecting the social networks' viewpoint. If there are fake, malicious or simply misinformed posts it's more likely that they are weakly renforced and through this cleanup their influence will be further diminished in the overall subsequent analysis.

Step 2: Determine existing linear scale

For the newly computed GAP< ~1000, the scale is a very good approximation of a linear function, and even for GAP< ~4500 we can find a linear trendline with a R^2 fit ≥0.99 (Figure 4). However, for larger values we need to employ different methods.

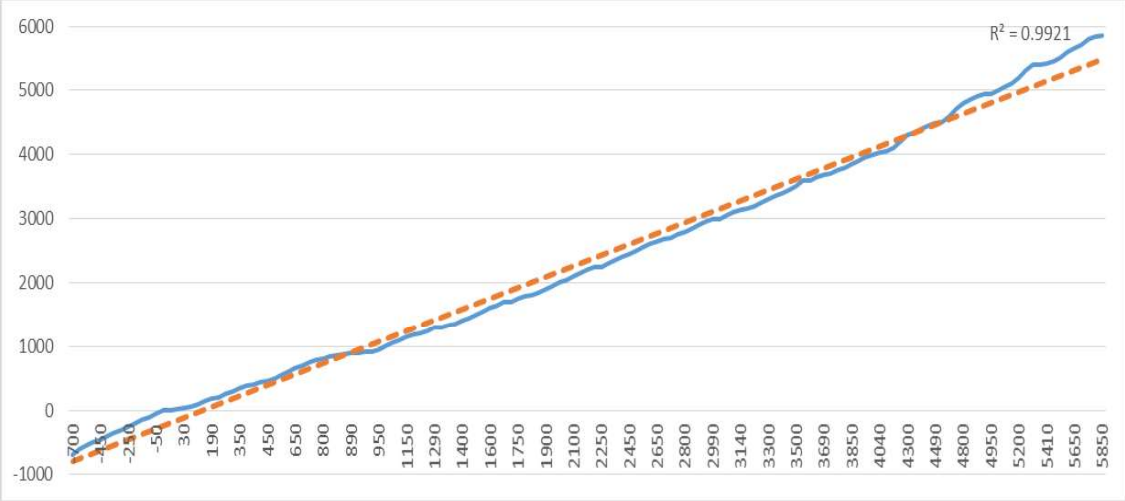

Figure 4: The quasi-linear part of GAP
The dashed line represents a linear trendline.

Step 3. Linearize the rest of the scale.

This was one of the most difficult parts, as there are few values and the scale is power-law. One purely mathematical method to solve this is to use a logarithmic value for GAP after a threshold. But, it is our oppinion that in order to model this part, we need to reverto the data intrinsec meaning: a validation of another's oppinion. So the question that arrises is: when is enough validation for an oppinion, i.e. : if 100 people agree that an oppinion is correct, does it matter that another 500 people agree with them? In order to answer this, we reffer to Table 2 and the fact that the average number of votes/post is 6, with the average GAP 3685. The simplest solution is to limit the maximum value at 4500, giving us a linear value scale and a limitative histogram with a number of 3423 positive oppinions (30%) limited at 4500 GAP. To further support this solution, it can be seen in Figure 5 that

there is a natural build-up in the histogram at values around 4500, the higher values frequencies being a less significant.

To summarize, in order to scale score with GAP, we currated the values, found the natural linear scale for GAP – partly due to the value variation and partly through limitation.
Using this linear GAP scale, OCS can be computed as shown in Figure 1, in order to be further used in various data analysis scenarios. The score itself is not really important in analyzing the social network's behaviour, more important is what the network members think and to what they react, which reflects in the GAP and votes variables.

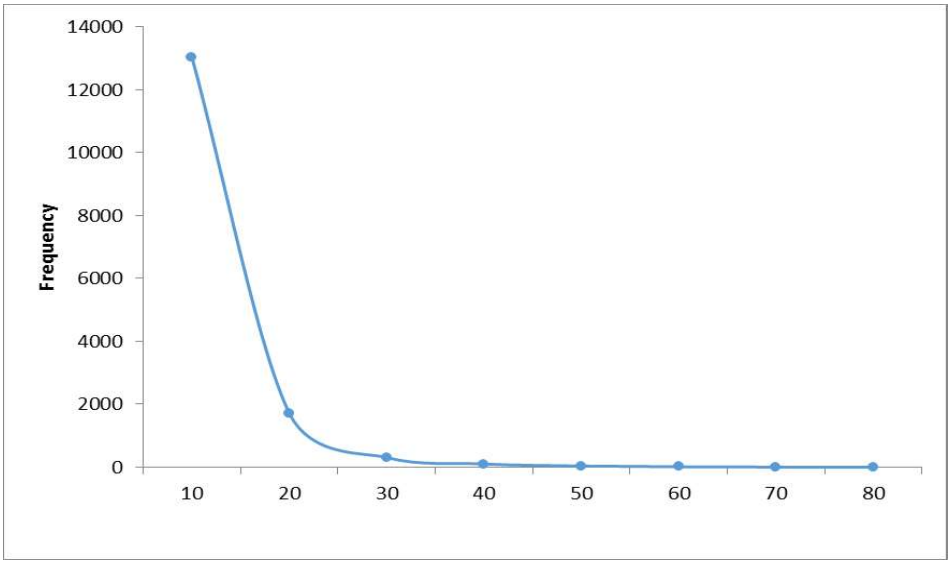

Figure 5: GAP histogram throughout the differrent transformation steps

Regarding the voter's behaviour the variables that express that are *GAP* and *votes* and they have a power-law distribution which is quite a clue (Figure 2,Figure 3b and Figure 6).In real social networks, one of the most used growth models is the Barabási-Albert which is based on the idea of preferential attachment (Barabasi, 2012). In this model, the network grows by adding new nodes that have the probability of attaching themselves to other nodes proportional to the number of links those nodes already have. The power-law degree distribution found in these types of networks is a consequence of this growth model.
In the case of the social network behind amfostacolo.ro, the aforementioned distributions of *GAP* and *score* are useful to validate that we are indeed working in a Barabási-Albert model, therefore hypothesis 1 "Social validation has a predictable behaviour in time" is true, as the validation obeys the growth laws of the model, the same as another network of this type.

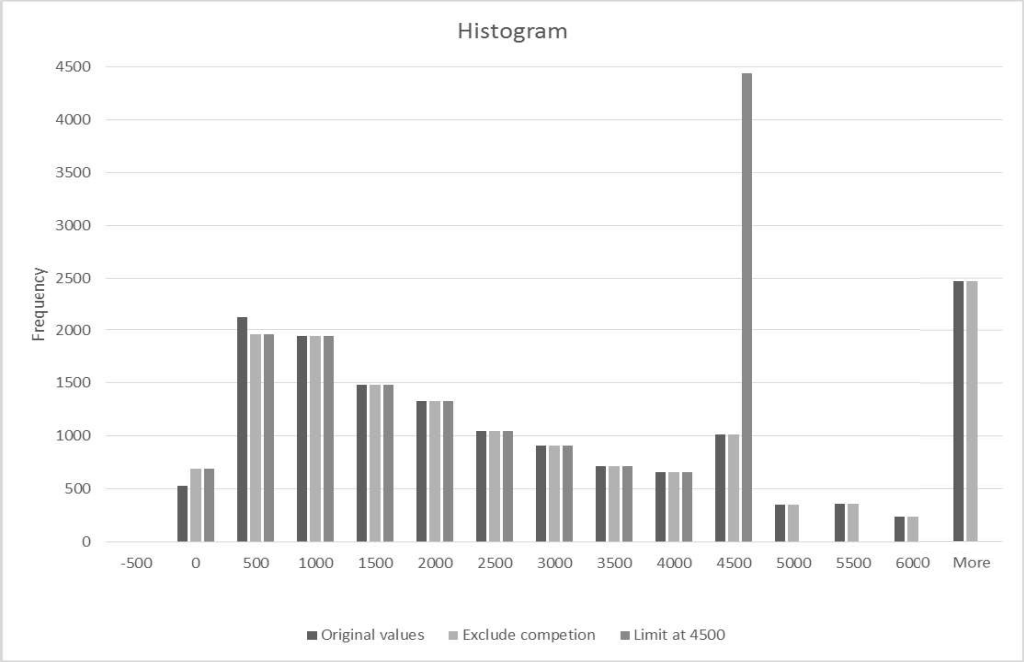

Figure 6: Number of votes (*votes* variable) histogram

The final hypothesis, H2: "On a large enough scale, spontaneous social aggregation selects the relevant information for the group" requires multifaceted understanding of the social network's behaviour that's reflected in the two social-influenced variables. Although there are many posts with few votes (13020 posts with ≤10 votes – 85.65% of the total) the remaining posts have an impressive number of votes and of GAPs. These are the socialy selected relevant informations (the ones that most people connect to) and this aggregation is especialy visible in large networks in which the underlining power-law distribution creates a dramatic difference (Albert-László Barabási, 2012). Otherwise formulated, people tend to vote and attribute trust to opinions that are already validated by many users and tend to abstain from voting on less-validated oppinions. In conclusion, for a majorly untempered, large review site, hypothesis 2 is true.

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 hotel managers. The problem that has arisen is way of validation of reviews,, posted on virtual platform such Amfostacolo.ro.

In this paper, we suggest way to determine the conditions under which relevant information arises from the comments posted.

The conclusions of the current study, carried out on the Am fost acolo.ro site have showed that:

5.1 On a large enough scale, spontaneous social aggregation selects the relevant informations for the group for a majorly untempered, large review site;

5.2 In the case of the social network behind amfostacolo.ro, the aforementioned distributions of GAP and score are useful to validate that we are indeed working in a Barabási-Albert model, therefore hypothesis “Social validation has a predictable behaviour in time” is true, as the validation obeys the growth laws of the model, the same as another network of this type.

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.

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

Bibliography

Ban, O.,Bădulescu, A. (2015) "Evaluarea online a calităţii serviciilor de cazare şi satisfacţia globală // The online evaluation of accommodation services quality and the overall satisfaction"*, Revista Româna 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.
Barabási, Albert-László (2012) *Network Science*, Ch 1-2. http://barabasilab.neu.edu/ networksciencebook/ Accesed in 12 December 2015
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.
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.

Ghosh, A., Varshney, S., Venugopal, P. (2014) "Social Media WOM: Definition, Consequences and Inter relationships", *Management and Labour Studies, August, 2014 39: 293-308.*
Hauser, J.R., Urban, G.L., & Weinberg, B.D. (1993) How consumers allocate their time when searching for information. Journal of Marketing Research, 30(4), 452–

466.
Mihalcea, R., Radev, D., (2011) *Graph-based Natural Language Processing and Information Retrieval,* Cambridge University Press ISBN 1139498827.
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.