Mining sentiments in Idea Management Systems as a tool for rating ideas

Adam Westerski
Universidad Politécnica de Madrid
Escuela Técnica Superior de Ingenieros de Telecomunicación
Avenida Complutense 30, Ciudad Universitaria, 28040 Madrid, Spain
westerski@dit.upm.es

Carlos A. Iglesias
Universidad Politécnica de Madrid
Escuela Técnica Superior de Ingenieros de Telecomunicación
Avenida Complutense 30, Ciudad Universitaria, 28040 Madrid, Spain
cif@gsi.dit.upm.es

ABSTRACT
In the contemporary Idea Management Systems one of the major challenges is rapid and automatic assessment of idea value. To address this problem, we propose the use of opinion mining technique for extracting sentiments from comments attached to ideas. Based on the opinion mining process, we introduce a new metric that summarises the community sentiment about an idea. We compare the performance of this metric with the currently used ones, as well as their impact on idea adoption. In particular, the study investigates behaviour of open-source communities based on data from Ubuntu Brainstrom - an Idea Management System instance run by Canonical to improve their Linux operating system distribution.

Keywords
idea management, information organization, opinion mining, sentiment analysis, metadata, ranking, portability

1. INTRODUCTION
One of the core characteristics of the Idea Management Systems is the participatory role of the community. The notion of crowd-sourcing is employed by inviting customers or employees to share and collaboratively improve their ideas. While this solution delivers valuable knowledge about the enterprise environment, it also introduces a number of problems related to the amount of data that needs to be processed by the idea competition organizers. In the contemporary solutions, the organization of large amounts of ideas is harnessed usually in two ways [8]: by rankings based on community metrics (top commented, top rated ideas etc.) or expert reviews (idea return of investment, client base impact etc.). In addition, the volume of data to assess is most often downsized by application of duplicate relationship for newly submitted ideas. Nevertheless, studies have shown [7] that those methods have a small impact on the final choice of ideas that are implemented and become very limited in instances that have collected tens of thousands of ideas.

As part our research done within the Gi2MO project [1] we sought to answer the question how to extend idea review capabilities and enable to characterize the semantics of innovation in a better way than with contemporary metrics. As reported in case studies [9], observing the reactions of clients on ideas is a very time consuming yet important element for the competition organizers in Idea Management Systems. Therefore, one of the solutions that we propose is analysis of the comments that users create when discussing the value of ideas and judging them.

In the following paper, rather then delivering a complete solution for this problem, we focus on analysing the relation between opinions mined from the idea comments and other automatically generated metrics (see Sec. 3). To do so, we introduce a new metric that aggregates the sentiment of comments attached to an idea and test our hypotheses using the dataset of Ubuntu BrainStorm (see Sec. 4). The results of our experiments (see Sec. 5) show that the newly introduced metric can be an interesting addition to Idea Management System and does not duplicate the contemporary metrics by delivering new information.

2. RELATED WORK
In the past years opinion mining has been a very active domain that vastly increased it's research activity [6, 16] along with the evolution of the Social Web and the growing popularity of Web 2.0 technologies [11]. The variety of approaches can be split into [10]: document subjectivity judgement, sentence analysis, or feature analysis. Depending on the taken approach, the contemporary solutions deliver accuracy ranging from 60 % for simple keyword methods though 80 % for various pattern matching or machine learning solutions [14], up to 90 % and above for domain optimized algorithms [17]. The tool used for research presented in the following paper treats idea comments as single documents and employs a keyword based approach.

In addition to development of different opinion mining approaches that improve the entire process accuracy, researchers have also proposed the inclusion of opinion mining into a pipeline of a larger scope. The usage of sentiment analysis has been evaluated in a number of domains such as: product review mining and summarization [18], business and government intelligence (e.g. trend prediction in sales [12]), analysis of public opinions before political elections [13].

Within the domain of Idea Management Systems in specific, there have been some attempts to employ opinion mining to improve idea review practices. In particular, Bothos et al. [5] proposed using opinion mining to improve prediction markets technique for rating ideas. While research conducted within Gi2MO project evaluated comparison of distributed

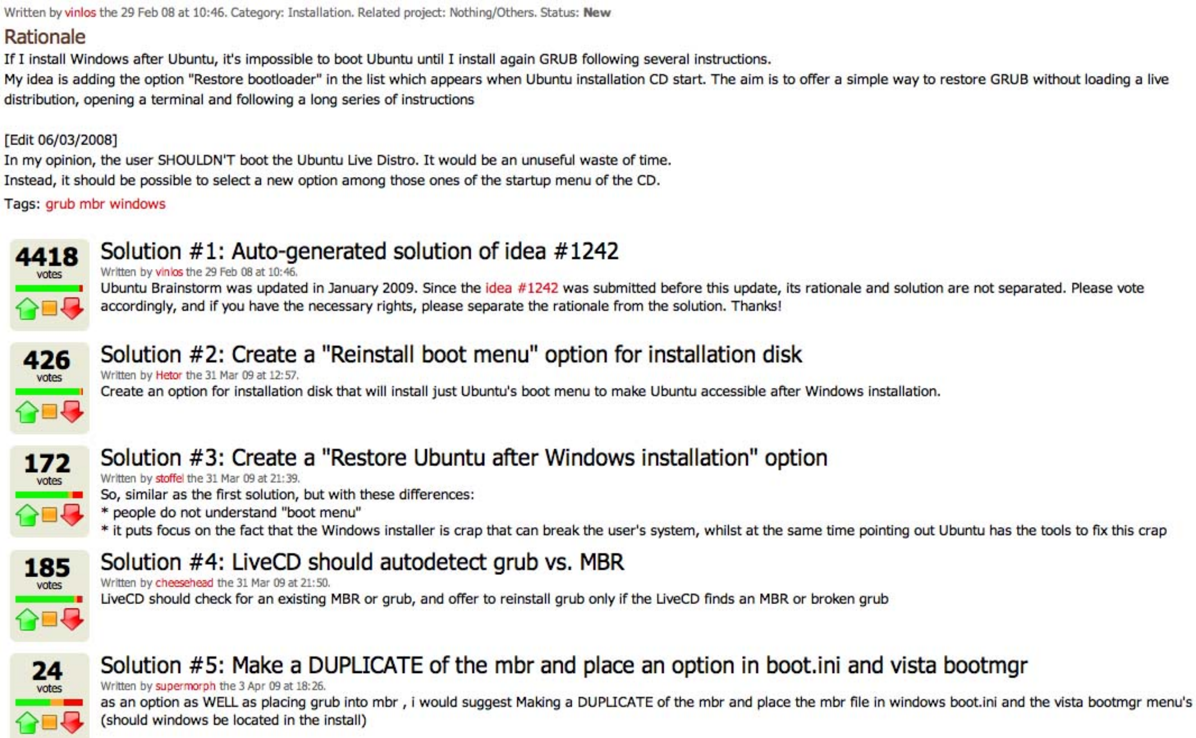

Figure 1: A sample single idea with solutions (Ubuntu Brainstorm [3]).

Idea Management Systems via sentiments of their communities [15]. In the following paper, we relate to both of those, however rather then focusing on details of application we peruse the evaluation of usefulness of textual opinions in Idea Management Systems in general. More precisely, we verify if inspecting community generated comments does actually influence the idea review workflow in a different way than other contemporary metrics. As such, our study aims to supplement previous work done in the area.

3. HYPOTHESIS

As shown in the previous section, the value of mining opinions from comments has been studied from many different angles and it's impact can differ depending on how the mined information is applied in practice. In the following paper we focus on two main hypotheses that relate opinion mining to Idea Management Systems:

H1. Organizations choose to implement ideas based on opinions of the community.

H2. Community opinions are not fully reflected by the currently used community activity metrics.

With **H1** we put forward a hypothesis that idea reviewers and managers of the idea competitions investigate not only the summery statistics like idea ratings but read the comments and those comments influence the final decisions that managers make in regard which ideas are implemented and which not.

With **H2** we suggest that the commonly used metrics in Idea Management Systems are not fully accurate about the opinions of the community regarding a certain idea. We hypothesise that evaluating opinions submitted in comments can deliver new knowledge that could potentially have additional impact on the final idea selections.

4. RESEARCH SETTING AND MEASURES

Taken into account both of the stated hypotheses, we propose to evaluate if they are indeed supported by evidence through calculating a single metric for every individual idea based on the following algorithm:

- calculate the opinion rating separately of every comment attached to the idea
- calculate the idea rating as a sum of ratings of it's comments

We applied the above methodology in practice using the dataset of Ubuntu Brainstorm [3]- an Idea Management System instance run by Canonical to collect ideas for improving their Ubuntu Linux distribution (see details in Table 1).

The distinctive feature of Ubuntu Brainstorm dataset, in comparison to other Idea Management data, is the possibility to submit new solutions for already existing ideas (see Fig. 1). The first solution is provided by the author of the idea, while the following solutions can be submitted by any member of the community. Each solution can be individually voted on, however the comments for all solutions are submitted in the same space, only referencing the root idea.

Table 1: Ubuntu Brainstorm dataset statistics

Metric	Metric Value
Idea number	21690
Comments number	133090
Users number	10062
Implemented Ideas number	576
Amount of Votes cast	2608917

In the preliminary work described in the following paper, in order to calculate the opinion rating per each comment, we constructed a simple prototype (OPAL [2]) that sums the word ratings of all words in the comment text (the word ratings were obtained using SentiWordNet library [4]). We measured the performance of such solution by manually annotating 50 idea comments (with positive, negative or neutral ratings) and compared the results with the automatic annotation done with OPAL. The proposed solution achieved 67 % recall, 66 % precision and 67 % f-measure.

Using the above method, we automatically annotated comments for 50 ideas: 10 implemented, 10 highest rated (with up/down rating), 10 lowest rated, 10 top commented, and 10 least commented (but having at least 1 comment). All together, we obtained opinion ratings for 1796 comments which were used to calculate the opinion ratings for the aforementioned 50 ideas.

Including the legacy metrics, we used the described dataset to calculate the following information:

- comment count - amount of comments attached to an idea
- solution count - amount of solutions submitted for an idea
- maximal solution up/down rating - the highest rating of a solution attached to an idea
- minimal solution up/down rating - the lowest rating of a solution attached to an idea
- average solution up/down rating - average of ratings of all solutions attached to an idea
- idea age - time (in days) since idea was submitted until the day experiment was conducted
- opinion rating - rating based on opinion mining algorithm run over comments attached to an idea
- idea adoption - indicates if an idea was implemented (equals 1) or not (equals 0).

To verify hypothesis **H1**, we analysed the impact of all legacy metrics on idea adoption (if an idea was implemented or not) and compared with the results for our opinion metric. To address hypothesis **H2** we analysed various correlations of our opinion metric with a number of currently utilised metrics in Idea Management Systems: community rating, comment count. etc. The results of those experiments are presented in the next section.

5. RESULTS

In case of hypothesis **H1**, for each of the aforementioned metrics we measured and compared the bivariate correlation with idea adoption to check if any of the metrics has a determining impact on whether ideas have been ultimately selected for implementation or rejected (see Table 2).

Table 2: Bivariate correlations of metrics with Idea Adoption

Metric name	Correlation
Comment count	0.03
Solution count	0.04
Max. solution rating	0.3
Min. solution rating	0.24
Avg. solution rating	0.37
Idea age	0.12
Opinion rating	**0.04**

The results show that correlation of opinion metric is one of the lowest. This suggests that reviewers and decision makers of the Ubuntu Brainstorm system did not pay attention to user opinions expressed in the comments. Such results indicate that hypothesis **H1** is not supported.

In the second activity to verify hypothesis **H2** we took the same metrics but measured the correlations between each other to see if opinion metric delivers new information or has the same behaviour as some other metric (see Table 3).

The obtained results show that opinion metric has a medium positive correlation with average rating, however weak correlation with max. rating and medium correlation with min. rating. Taking into account this result, we can make a statement that in the particular settings of Ubuntu Brainstorm good ratings of idea solutions do not reflect the community opinions, while poor ratings usually go in line with bad comments. To confirm this observation, we also investigated the raw data of the max. and min. solution rating metrics. Figure 2 shows that the behaviour of min. rating is similar to opinion rating in the area of solutions with lowest rating (2), while in other areas the similarities are much harder to observe (especially in top voted area (1) where some of the top ideas have lowest opinion rating of the entire sample).

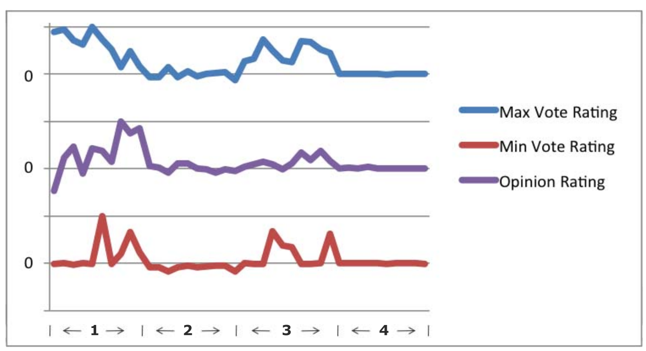

Figure 2: Comparison of normalized rating values for ideas from the experiment. (1) - top voted solutions, (2) - lowest voted, (3) - implemented, (4) - least commented

Table 3: Bivariate correlations of metrics with each other (including opinion rating).

		1	**2**	**3**	**4**	**5**	**6**	**7**
1	Comment count	1	0.37	0.68	0.11	0.26	0.02	**0.28**
2	Solution count	x	1	0.28	-0.32	-0.21	-0.65	**-0.08**
3	Max. solution rating	x	x	1	0.32	0.51	0.32	**0.25**
4	Min. solution rating	x	x	x	1	0.95	0.26	**0.38**
5	Avg. solution rating	x	x	x	x	1	0.26	**0.41**
6	Idea age	x	x	x	x	x	1	**0.19**
7	Opinion rating	x	x	x	x	x	x	1

Moreover, this criticism of most down ranked ideas or lack of support for top ranked ideas should not be understood in terms of quantity of opinions (due to weak correlation between opinion rating and comment count) but strength and verboseness of sentiment expressions in the comments. Taking into account those results and observing the correlation of opinion rating metric with the remaining legacy metrics we can conclude that the new metric does not duplicate the behaviour of other Idea Management indicators. Therefore, hypothesis **H2** can be considered as supported.

6. FUTURE WORK

The presented results are a preliminary work that was done to determine the perspectives for perusing opinion mining topic in Idea Management Systems and to compliment our reach in Gi2MO project on modelling the knowledge of opinions during community discussions over the Web [15]. While, the results of experiments presented here are promising, in terms of future work we intend to investigate the topic further and support our study with better evidence. Firstly, we envision repeating the described experiment with manual annotation that would eliminate the uncertainty that comes from using an opinion mining algorithm with a fairly low performance, like the OPAL prototype constructed by the Gi2MO team. Another option for future improvement that we would like to point out is using a better opinion mining algorithm, aligned to the domain of Ubuntu Linux and presenting similar evaluation but for the entire dataset of 21000 ideas. In addition, the creditability of the presented results could be increased by presenting a study for a number of different Idea Management instances of different domains and vendors.

7. CONCLUSIONS

Concluding our investigation we can state that measuring community sentiments related to ideas through comments does deliver a supplementary tool for judgement of ideas performance, however not fully distinct from the idea rating metric. The strongest detected correlation with average solution rating suggests that the up/down rating does partially reflect what users write in the comments. The interesting observation of our analysis is that very well rated solutions of ideas often attract high amount of criticism as well as positive feedback (low correlations of the top and lowest ratings with opinion rating).

In addition, our study has shown that the impact of user opinions is very low on final idea selections, proving that idea reviewers and contest managers responsible for selecting ideas to implement do not make their choice based on opinion of the community.

8. ACKNOWLEDGMENTS

This research has been partly funded by the Spanish Ministry of Industry, Tourism and Trade through the project RESULTA (TSI-020301-2009-31) and Spanish CENIT project THOFU. We express our gratitude to Atos Origin R&D for their support and assistance.

9. REFERENCES

[1] Gi2mo project homepage. `http://www.gi2mo.org/`, 2012.

[2] Opal (opinion analyser) application homepage. `http://www.gi2mo.org/apps/opal-opinion-analyser/`, 2012.

[3] Ubuntu brainstorm idea management instance. `http://brainstorm.ubuntu.com/`, 2012.

[4] S.~Baccianella, A.~Esuli, and F.~Sebastiani. Sentiwordnet 3.0: An enhanced lexical resource for sentiment analysis and opinion mining. In *Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC'10)*, Valletta, Malta, May 2010. European Language Resources Association (ELRA).

[5] E.~Bothos, D.~Apostolou, and G.~Mentzas. Idem: A prediction market for idea management. designing e-business systems. markets, services, and networks. In *7th Workshop on E-Business, WeB 2008*, Paris, France, 2008.

[6] A.~Esuli. Sentiment classification bibliography. `http://liinwww.ira.uka.de/bibliography/Misc/Sentiment.html`, 2007.

[7] P.~M.~D. Gangi and M.~Wasko. Steal my idea! organizational adoption of user innovations from a user innovation community: A case study of dell ideastorm. *Decision Support Systems*, 48:303– 312, 2009.

[8] S.~Hrastinski, N.~Z. Kviselius, and M.~Edenius. A review of technologies for open innovaton: Characteristics and future trends. In *Proceedings of the 43rd Hawaii International Conference on System Sceiences*, 2010.

[9] G.~Jouret. Inside cisco's search for the next big idea. *Harvard Business Review*, September 2009.

[10] B.~Liu. Opinion mining & summarization - sentiment analysis. Tutorial at International World Wide Web Conference 2008 (WWW2008), Beijing, China, April 2008.

[11] M.~D. Lytras and P.~O. de~Pablos. *Social Web Evolution: Integrating Semantic Applications and Web 2.0 Technologies.* Information Science Reference, 2009.

[12] G.~Mishne and N.~Glance. Predicting movie sales from blogger sentiment. In *AAAI Symposium on*

Computational Approaches to Analysing Weblogs (AAAI-CAAW), New York, NY, USA, 2006.

[13] T.~Mullen and R.~Malouf. Taking sides: User classification for informal online political discourse. *Internet Research*, 18:177– 190, 2008.

[14] B.~Pang, L.~Lee, and S.~Vaithyanathan. Thumbs up? sentiment classification using machine learning techniques. In *Proceedings of the 2002 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 79–86, Philadelphia, PA, USA, May 2002.

[15] A.~Westerski, C.~A. Iglesias, and F.~T. Rico. Linked opinions: Describing sentiments on the structured web of data. In *4th international workshop Social Data on the Web (SDoW2011)*, Bonn, Germany, October 2011.

[16] J.~Wiebe. Opinion mining bibliography by jan wiebe. `www.cs.pitt.edu/~wiebe/subjectivityBib.html`, 2012.

[17] H.~Yu and V.~Hatzivassiloglou. Towards answering opinion questions: Separating facts from opinions and identifying the polarity of opinion sentences. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 2003.

[18] L.~Zhuang, F.~Jing, and X.-Y. Zhu. Movie review mining and summarization. In *Proceedings of the 15th ACM international conference on Information and knowledge management (CIKM '06)*, New York, NY, USA, 2006.