	Evangelists				Detractors			
	Precision		Recall		Precision		Recall	
Rank size	IT	FF	IT	FF	IT	FF	IT	FF
TOP 10	10.0%	10.0%	10.0%	10.0%	40.0%	40.0%	57.0%	57.0%
TOP 20	25.0%	15.0%	50.0%	30.0%	30.0%	30.0%	85.0%	85.0%
TOP 30	20.0%	16.7%	60.0%	50.0%	20.0%	20.0%	85.0%	85.0%
TOP 40	15.0%	15.0%	60.0%	60.0%	15.0%	15.0%	85.0%	85.0%
TOP 50	14.0%	12.0%	70.0%	60.0%	14.0%	12.0%	100.0%	85.0%
TOP 60	15.0%	10.0%	90.0%	60.0%	11.6%	10.0%	100.0%	85.0%
TOP 70	12.9%	8.6%	90.0%	60.0%	10.0%	8.6%	100.0%	85.0%
TOP 80	11.3%	7.5%	90.0%	60.0%	8.8%	7.5%	100.0%	85.0%
TOP 90	10.0%	6.7%	90.0%	60.0%	7.8%	6.7%	100.0%	85.0%
TOP 100	9.0%	6.0%	90.0%	60.0%	7.0%	6.0%	100.0%	85.0%
TOP 110	8.2%	5.5%	90.0%	60.0%	6.4%	5.5%	100.0%	85.0%
TOP 120	7.5%	5.8%	90.0%	70.0%	5.8%	5.0%	100.0%	85.0%
TOP 130	6.9%	5.4%	90.0%	70.0%	5.4%	4.6%	100.0%	85.0%
TOP 140	7.1%	5.0%	100.0%	70.0%	5.0%	4.3%	100.0%	85.0%

Table 4: Precision and Recall (Evangelists)

This happens due to the occasional difficulty for distinguishing between a neutral and a positive-biased tweet during the manual classification. For the negative tweets, this boundary is usually clearer.

Since there is no benchmark for influential users detection (a default dataset with tweets and users previously classified), one significant effort of this work was to built such a test collection. This is not a trivial task due to the difficulty to classify posts as positive or neutral (this is a subjective problem by nature).

The experiments results also demonstrate that the interactions (mentions, replies, re-tweets, attributions) of an user with others is a better representation of her influence than her connections (follower, following). The precision and recall values for the generated ranks, using the interactions, were always better. Another substantial remark is that the interaction network is more sparse than the relations network. This turns the computational cost must cheaper and with more accurate results.

As future work, we can address the problem of finding α, β and γ parameters. The developed technique also needs further testing in real environments (with evangelists and detractors identified) and on data bases with different themes.

Acknowledgments. The authors would like to thank Rodrigo Pazzini for the analysis of user influence and insightful comments. This work is partially supported by the projects INCT-Web (MCT/CNPq grant 57.3871/2008-6) and by the authors' individual grants and scholarships from CNPq, CAPES and FAPEMIG.

7. REFERENCES

[1] R. A. Baeza-Yates and B. Ribeiro-Neto. *Modern Information Retrieval*. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 1999.

[2] P. Bonacich. Some unique properties of eigenvector centrality. *Social Networks*, 29(4):555 – 564, 2007.

[3] M. Cha et. al. Measuring User Influence in Twitter: The Million Follower Fallacy. In *Procs. Intl. AAAI Conf. on Weblogs and Social Media (ICWSM)*.

[4] L. F. Costa et. al. Characterization of complex networks: A survey of measurements. *Advances In Physics*, 56:167, 2007.

[5] D. H. Dalip et.al. Automatic quality assessment of content created collaboratively by web communities: a case study of wikipedia. In *Procs. Joint Conf. on Digital libraries (JCDL)*, pages 295–304, 2009.

[6] N. A. Diakopoulos and D. A. Shamma. Characterizing debate performance via aggregated twitter sentiment. In *Procs. Intl. Conf. on Human Factors in Computing Systems (CHI)*, 2010.

[7] B. A. Huberman, D. M. Romero, and F. Wu. Social networks that matter: Twitter under the microscope. *ArXiv e-prints*, December 2008.

[8] A. Java et. al. Why we twitter: understanding microblogging usage and communities. In *Procs. of WebKDD and SNA-KDD Workshop on Web Mining and Social Network Analysis*, pages 56–65, 2007.

[9] B. Krishnamurthy, P. Gill, and M. Arlitt. A few chirps about twitter. In *Procs. Workshop on Online Social Networks (WOSP)*, pages 19–24, 2008.

[10] A. Leavitt et. al. New approaches for analyzing influence on twitter. Technical report.

[11] M. E. J. Newman. Power laws, pareto distributions and zipf's law. *Contemporary Physics*, 46(5), 2005.

[12] S. Ressler. *Perspectives on electronic publishing: standards, solutions, and more.* Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 1993.

[13] T. Rodrigues, F. Benevenuto, V. Almeida, J. Almeida, and M. Gonçalves. Uma análise contextual de conteúdo duplicado no youtube. In *WebMedia 2009*, Fortaleza, CE, Brasil, 2009.

[14] Twitter Blog. Measuring Tweets. `http://blog.twitter.com/2010/02/measuring-tweets.html`, 2010.

[15] J. Weng et. al. Twitterrank: finding topic-sensitive influential twitterers. In *Procs. ACM Intl. Conf. on Web Search and Data Mining (WSDM)*, pages 261–270, 2010.