Reputational DAMAGE: Classifying the impact of allegations of irresponsible corporate behavior expressed in the financial media

Andy Moniz[1] and Franciska de Jong[2,3]

[1]Rotterdam School of Management, Rotterdam, The Netherlands
{moniz}@rsm.nl
[2] Erasmus Studio, Erasmus University, Rotterdam, The Netherlands
{f.m.g.dejong}@eshcc.eur.nl
[3] Human Media Interaction, University of Twente, Enschede, The Netherlands
{f.m.g.dejong}@utwente.nl

Abstract

In this study we design an automated Online Reputation Management (ORM) system that mines news texts by comparing the financial media's attributions of culpability during a corporate reputational crisis, with those expressed by the firm itself. Our working assumption is that the stronger the media's attributions of organizational responsibility, the more likely it is that the attributions will influence public perceptions and damage a firm's reputation. The system presented here works in four steps: the first phase employs a multinomial Naïve Bayesian model that detects irresponsible corporate behavior (e.g. bribery, fraud, negligence). The second phase employs a Latent Dirichlet Allocation (LDA) topic model to infer attributions of corporate culpability during a crisis. The third stage computes a measure of document polarity by counting terms from the General Inquirer dictionary. Finally, the components are combined into an ensemble tree that classifies the likelihood that a given media allegation may damage corporate reputation. Our findings suggest that human perceptions intensify during corporate crises particularly if a firm exhibits signs of arrogance and denial, while reputational damage may be mitigated if a firm is seen to address the concerns expressed by the media.

1. Introduction

Prior text mining studies of corporate reputational sentiment have mostly captured reputation by classifying the impact of consumer sentiment from product reviews and tweets on corporate brands [1]. The goal of this study is to provide an insight into human perceptions by examining the formation of attributions during corporate crises. Prior research suggests that, given a corporate crisis, a firm should design an external communication strategy aimed at mitigating the impact of reputational damage by altering public perceptions [2,3]. Such strategies range from defensive communications that place the firm's interests first (e.g. denial), to accommodative communications that put victims' concerns first. The greater the crisis responsibility generated by the crisis, the more accommodative the strategies should be.

To date, research into crisis communications strategies has considered only specific crisis types (e.g. product recalls), thereby limiting the ability to generalize the findings for human behavior and to make recommendations for corporate crisis communications' departments. In this study, we consider a range of media allegations of irresponsible corporate behavior and classify the types of allegations that pose the greatest reputational threat to a firm. Our study provides an insight into human behavior by comparing media attributions of corporate culpability with the views expressed by firms in their public communications. From an applications perspective, our results may be of interest to corporates seeking to design and/or integrate an Online Reputation Management (ORM) system to develop more effective corporate communications in order to mitigate corporate reputational damage.

The rest of this paper is structured as follows. Section 2 draws on literature from the field of organizational studies, by discussing corporate crisis communication strategies, and the implications for sentiment and reputational analysis. Section 3 explains the individual components of the ORM system. In

Section 4, outlines the financial media corpus and survey-based reputation scores used for evaluation. We provide an evaluation of the ORM components and discuss the results. Section 5 concludes and suggests avenues for future research.

2. Related Literature

2.1 Background: crisis communications

At the core of our proposed model lies Situational Crisis Communication Theory (SCCT), developed in the field of organizational studies [4]. Crisis responsibility, the degree to which the public attributes responsibility for a crisis to an organization, is the centerpiece of SCCT [4] and states that public attributions of crisis responsibility are directly related to the reputational threat posed by a crisis.

The implication of SCCT is that similar types of crises can be managed in similar ways. For example, the 'victim' cluster contains crisis types that produce very low attributions of crisis responsibility (e.g. natural disasters) and represent a mild reputational threat. Organizations are viewed as victims of the crisis because the crises are seen as driven by external forces that were beyond management's control [5]. The 'accidental' cluster produces minimal attributions of crisis responsibility (e.g. technical errors) and represent a moderate reputational threat. The public may believe that the firm's management did not intend the crisis to happen and/or could do little to prevent it [5]. Finally, the 'intentional' crisis cluster contains crisis types that produce strong attributions of crisis responsibility (e.g. irresponsible corporate behavior) and represent a severe reputational threat, as corporate management knowingly violated laws and/or placed the public and employees at risk.

The outcome of SCCT is the recommendation that crisis managers should deny responsibility in the case of rumors, apologize for accidents, and undertake corrective action for intentional, irresponsible behavior [5]. We draw on this theory by developing an ORM system that is intended to infer media attributions of corporate culpability following allegations of irresponsible corporate behavior, and provides a monitoring tool for corporations to develop a more effective communications response to mitigate reputational crises.

2.2 Reputational polarity analysis

Mining and interpreting opinions about companies is a harder and less understood problem than opinion mining for products and services. Firm reputation is an intangible metric [6] and may be viewed differently by different stakeholder groups (e.g. consumers, investors, regulators, and local communities), making reputation analysis a challenging task. Stakeholder groups may weigh criteria differently when evaluating the reputation of a firm, hindering the ability to systematically classify news text into one with a positive or negative sentiment without first defining the stakeholder perspective. For example, a regulatory imposed fine for an oil spill may be seen positively from the point of view of the public (due to a sense of justice) and negatively from the stance of investors (due to the penalty reducing corporate earnings). Our proposed approach measures corporate reputation from the perspective of investors. As ultimate owners of a firm, investors are arguably one of the most important stakeholder groups for a firm. To this end, we use a financial media news source to retrieve allegations of irresponsible corporate behavior, and use survey-based reputational rankings from financial analysts and corporate executives.

3. Model of Reputational Damage

In this section we describe the four components of the proposed ORM system. The first phase employs a term counting approach to detect irresponsible corporate behavior, while the second phase implements a topic model to infer media attributions of corporate culpability. The third step measures document polarity and the final stage employs an ensemble tree to combine the three components.

3.1 Event detection model

The first phase of the system detects irresponsible behavior by employing a multinomial Naïve Bayesian model. We label the approach the event detection model. Following prior text mining studies in the field of finance our main resource to identify document terms is the General Inquirer dictionary [7]. The dictionary contains 1,915 positive words and 2,291 negative words. Negative terms include: 'accident', 'illegal' and 'negligence'. We perform a pre-processing step that consists of Snowball stemming and stop word removal and select the 1500 most frequent words in the training set (described in section 4.1) selected using a simple binary weighting scheme. Previous research has found the weighting scheme achieves higher accuracy for sentiment analysis than term frequency weighting [8]. For illustrative purposes, Figure 1 displays the resulting clustering of terms calculated using the TextRank algorithm [9].

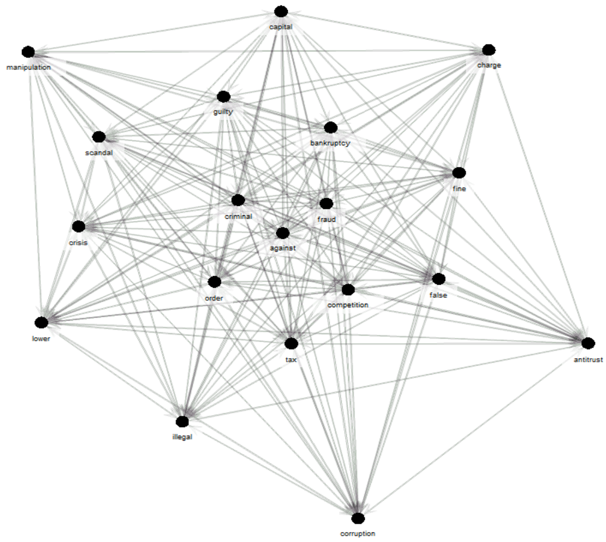

Fig. 1. Link analysis of frequently occurring negative terms

3.2 Attribution topic model

The second phase of the system extends the baseline term-counting method by setting attribution priors in a Latent Dirichlet Allocation (LDA) topic model [10]. A traditional topic, in topic models such as LDA, is a list of words sorted by the probability that each word can be generated from the topic. In the attribution-topic model, a topic is related to the distribution of media attributions of corporate culpability. Our approach is motivated by [11] and draws on the concept of 'intensifers' [12,13] that magnify the degree of expressed sentiment in a document. Our working assumption is that the stronger the media's attributions of organizational responsibility, the more likely it is that a crisis will influence investors' perceptions and damage on firm reputation [5].

We employ attribution topics as priors in the LDA clusters to capture global attribution topics in documents [14] inferred from the financial media. We implement standard settings for LDA hyperparameters [15] with $\alpha = 50/K$ and β=.01, where K is the number of topics. To be consistent with the heuristic approach adopted by prior organizational studies [16] we set K equal to five. Table 1 identifies the top terms associated with each of the clusters of the ORM model. Representative words are the highest probability document terms for each topic cluster. The inferred aspects are manual annotations associated with the topic clusters.

Table 1. Topic clusters and top words identified by LDA

unintentional		corporate scandal		corporate chastise		abhorrence		corporate arrogance	
word	**prob.**	**word**	**prob.**	**word**	**prob.**	**word**	**prob.**	**word**	**prob.**
murky	0.100	scandalous	0.117	chastise	0.221	severe	0.193	arrogance	0.149
problem	0.099	restless	0.056	cringe	0.135	complicate	0.077	audacious	0.145
accuse	0.071	ghastly	0.048	scold	0.130	liable	0.069	accuse	0.114
protest	0.048	cynical	0.044	altercation	0.101	infuriate	0.064	foreboding	0.112
terrorism	0.046	critic	0.037	falsehood	0.061	explode	0.053	warlike	0.107
accident	0.041	ludicrous	0.036	fraught	0.040	mortify	0.052	rebellion	0.090

We refer to the topic probabilities associated with each topic cluster as the attribution topic model.

3.3 Polarity detection

The third phase of the ORM measures document polarity by counting the number of positive (P) versus negative (N) terms using the General Inquirer dictionary [20]. The approach is consistent with the methodology adopted in the field of finance [17] to measure financial media sentiment associated with a company's stock market patterns. Our measure of document polarity is included as a third component in the ensemble tree.

4. Experiments

In this section we discuss the corpus of financial media allegations and describe the Fortune reputational survey ratings that are used to evaluate the ORM system. We then evaluate the ensemble classification tree, present the results and provide a discussion.

4.1 Data

Our news source is a corpus created from Dow Jones Newswires (DJNW), and is commonly used within financial literature [18]. News articles are sourced from financial blogs, (e.g. MarketWatch.com) and the on-line editions of financial newspapers (e.g. The Wall Street Journal). We include the 'Editorial Commentary' and 'Letters to the Editor' sections of newspapers on the assumption that these articles contain more opinionated views than fact based articles [19]. We separately source corporate press releases related to the media allegations from PR Newswires. We conduct keyword searches on the headline and the first sentence of news stories that match the terms 'accusation' or 'allegation' in lemmatized form. These words were chosen because they convey negative connotations of irresponsible corporate behavior, even though they are insufficient in their own right to determine the nature, severity and cause of an incident for a potential reader to determine the potential impact on corporate reputation. We conduct multiple name searches using variants of companies' names obtained from companies' websites, Wikipedia and the Open Directory Project (ODP) since firms are often referred to by their popular names rather than legal names (e.g.'IBM' rather than 'International Business Machines Inc'). Our resulting corpus consists of 35,678 news stories for 598 unique global companies during the period January 2009 to December 2013. For the purposes of training the event detection model in Section 3.1, we separately collected media allegations over the period 2006-2009.

For evaluation we obtain reputational ratings from Fortune magazine's list of the World's Most Admired Companies, one of the most prominently used proxies of investors' perceptions of corporate reputation [20]. Fortune surveys approximately 15,000 executives and financial analysts to identify annual changes in corporate reputation. To classify the likelihood of reputational damage, we create a binary variable that equals one if the company's reputation declines over the course of one year, and zero otherwise. While our decision to evaluate daily media allegations against an annual reputation score is not ideal, we note that it is the standard approach used in the fields of organizational studies [32] and finance [30] given the absence of a daily reputation score. In subsequent research, we intend to address this concern by employing Amazon's Mechanical Turk.

4.2 Experiment setup

Learning and prediction is performed using an ensemble classification tree. In line with [8], our goal is not to design a system that outperforms state-of-the-art machine learning methods, but to identify an approach that corporate communications departments can adopt by following a set of more transparent rules and thresholds. We adopt the Random Forest algorithm [21] which uses a diverse set of classifiers by introducing randomness into the classifier construction.

Experiments were validated using 10-fold cross validation; the dataset is broken into 10 equal sized sets, the classifier is trained on 9 datasets and tested on the remaining dataset. The process is repeated 10 times and we calculate the average across folds. To evaluate model classification, we select precision and recall measures defined as:

$$Precision = \frac{\#\ correct\ predictions\ of\ declining\ reputation}{total\ \#\ of\ predictions\ of\ declining\ reputation}$$

$$Recall = \frac{\#\ correct\ predictions\ of\ declining\ reputation}{total\ \#\ of\ declines\ in\ corporate\ reputation}$$

4.3 Experiment results and discussion

To enhance our understanding of the ORM system we separately evaluate the event detection and attribution components. The components are included in the ensemble tree together with the polarity measure. The evaluation metrics are shown in Table 2.

Table 2: Model evaluation

Model	Precision	Recall
event detection model	0.690	0.309
attribution topic model	0.277	0.131
joint model	0.875	0.217

Our results suggest that the event detection model and attribution topic model capture distinct dimensions of media sentiment. Classifying the impact of allegations on corporate reputation requires both the detection of the initial triggering event, and the corresponding media attribution of corporate culpability. As expected recall measures are relatively low, as daily media allegations are evaluated against an annual reputation score.

To aid our understanding of the ORM system, Figure 2 displays the decision tree results for one of the folds. The grey boxes provide the underlying probabilities associated with the classification of declining reputation. A value of 1.0 implies there is a 100% likelihood that the media allegation will damage firm reputation.

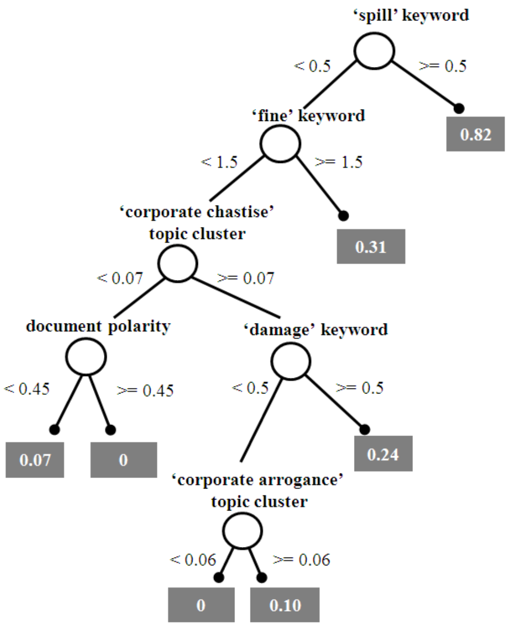

Fig 2: Example classification tree from one fold

The ORM system identifies that environmental contamination spills and government/regulatory fines are most likely to damage corporate reputation. The next important decision in the tree is the degree to which the financial media 'chastises' the company for its behavior, as captured by the attribution topic model. This finding is consistent with organizational studies that suggests that stronger attributions of corporate culpability represent a more severe reputational threat [5]. Finally, we highlight the interaction between the 'corporate chastise' topic attribution cluster and the sentiment expressed in a corporation's communication response to irresponsible corporate behavior. Intriguingly, reputational damage increases if the corporation exhibits signs of arrogance (e.g. by denying the crisis), yet if a company issues a press release without displaying arrogance (e.g. potentially by offering a full apology), there is no lasting impact on firm reputation.

5. Conclusion

In this paper we propose to model sentiment analysis for online new reports on irresponsible corporate behavior and to integrate the classification obtained in a corporate reputation management system through a decision tree. Our approach compares the financial media's attributions of culpability during a corporate reputational crisis with those expressed by the firm itself.

Our findings suggest that the willingness of a corporation to accept responsibility for a crisis may help mitigate the impact of reputational damage. In future research the impact of reports from other stakeholder groups (ranging from consumers, employees and special interest groups), expressed via social media and other online sources, and the linguistic features of corporate communications will be integrated into our study in a more in-depth way.

Acknowledgement

The research leading to these results has partially been supported by the Dutch national program COMMIT.

References

[1] Villena-Roma, J., Lana-Serrano, S., Moreon-Garcia, C., Garcia-Moera, J., Gonzalez-Cristobal, J.C., 2012. RepLab 2012: Polarity Classification and Filtering on Tweet Data.

[2] Coombs, W., T. 1999b. Ongoing crisis communication: Planning, managing, and responding. Sage

[3] Marcus, A. A., and Goodman, R. S., 1991. Victims and shareholders: The dilemmas of presenting corporate policy during a crisis. Academy of Management Journal, 34

[4] Coombes, W., T. 2007. Protecting Organization Reputations During a Crisis: The Development and Application of Situational Crisis Communication Theory. Corporate Reputation Review,Vol. 10, No. 3, pp. 163–176

[5] Coombs , W . T . and Holladay , S . J . 2005. Exploratory study of stakeholder emotions: Affect and crisis. N.M. Ashkanasy, W.J. Zerbe and C.E.J. Hartel (eds.)

[6] Barney, J. 1991. Firm resources and sustained competitive advantage. Journal of Management, 17: 99–120.

[7] Stone, P., Dumphy, D. C., Smith, M. S., and Ogilvie, D. M., 1966. The General Inquirer: A Computer Approach to Content Analysis. The MIT Press.

[8] Pang, B., Lee, L., and Vaithyanathan, S. 2002. Thumbs up? Sentiment classification using machine learning techniques. Proceedings of EMNLP-02.

[9] Mihalcea, R., and Tarau, P. TextRank: Bringing order into texts. In Proceedings of the 2004 Conference on Empirical Methods in Natural Language Processing

[10] Blei, D., M., Ng, A., Jordan, M., I., 2003. Latent Dirichlet Allocation, Journal of Machine Learning Research 3, 993-1022

[11] Bao , S., Xu, S., Zhang, L., Yan, R., Su, Z., Han, D., Yu, Y. 2009. Joint Emotion-Topic Modeling for Social Affective Text Mining. Ninth IEEE International Conference on Data Mining

[12] Kennedy, A., Inkpen, D., 2006. Sentiment Classification of Movie Reviews using Contextual Valence Shifters. Computational Intelligence, vol.22(2), pp.110-125, 2006

[13] Polanyi, L., Zaenen, A., 2004. Contextual Valence Shifters, AAAI Spring Symposium on Exploring Attitude and Affect in Text: Theories and Applications

[14] Titov, I, McDonald, R.T., 2008. Modeling online reviews with multi-grain topic models. Proceedings of the 17th international conference on World Wide Web, 111-120, 253

[15] Griffiths, T. L., Steyvers, M., 2004. Finding scientific topics. Proceedings of the National Academy of Science, 101, 5228-5235

[16] Murphy, D., Shrieves, R., Tibbs, S., 2009. Understanding the Penalties Associated with Corporate Misconduct: An Empirical Examination of Earnings and Risk. Journal of Financial and Quantitative Analysis

[17] Tetlock, P., Saar-Tsechansky, M. Macskassy, S., 2008. More Than Words: Quantifying Language to Measure Firms' Fundamentals. Journal of Finance, Vol. LXIII, No. 3, 1437-1467

[18] Tetlock, P., 2007. Giving Content to Investor Sentiment: The Role of Media in the Stock Market. The Journal of Finance 62, no. 3, (June 1): 1139.

[19] Kozareva, Z., B. Navarro, S. Vazquez, and A. Nibtoyo.2007.UA-ZBSA: A Headline Emotion Classification through Web Information. Proceedings of the 4th International Workshop on Semantic Evaluations.

[20] Wartick, S., L. 1992. The relationship between intense media exposure and change in corporate reputation. Business & Society, 31(1), 33-49.

[21] Breiman, L., 2001. Random Forests. Machine Learning