FAKE REVIEW ANNOTATION MODEL AND CLASSIFICATION THROUGH REVIEWERS' WRITING STYLE

SOMAYEH SHOJAEE

FSKTM 2020 3

FAKE REVIEW ANNOTATION MODEL AND CLASSIFICATION THROUGH REVIEWERS' WRITING STYLE

By

SOMAYEH SHOJAEE

Thesis Submitted to the School of Graduate Studies, Universiti Putra Malaysia, in Fulfilment of the Requirements for the Degree of Doctor of Philosophy

September 2019

COPYRIGHT

DEDICATIONS

To Hannah and Saeid

Abstract of thesis presented to the Senate of Universiti Putra Malaysia in fulfilment of the requirement for the degree of Doctor of Philosophy

FAKE REVIEW ANNOTATION MODEL AND CLASSIFICATION THROUGH REVIEWERS' WRITING STYLE

By

SOMAYEH SHOJAEE

September 2019

Chair: Assoc. Prof. Masrah Azrifah Azmi Murad, PhD
Faculty: Computer Science and Information Technology

In the last decade, online product reviews have become the main source of information during customers' decision making and business' purchasing processes. Unfortunately, fraudsters have produced untruthful reviews driven intentionally for profit or publicity. Their activities deceive potential organizations to reshape their businesses, customers from making best decisions and opinion mining techniques from reaching accurate conclusions.

One of the big challenges of spam review detection is the lack of available labeled gold standard real-life product review dataset. Manually labeling product reviews as fake or real is one of the approaches to deal with the problem. However, recognizing whether a review is fake or real is very difficult by only reading the content of the review, because spammers can easily craft a fake review that is just like any other real reviews.

To address this problem we enhance the inter-annotator agreement in manually labeling approach by proposing a model to annotate product reviews as fake or real. This is the first contribution of this research study. The proposed annotation model is designed, implemented and accessed online. Our crawled reviews are labeled by three annotators who were trained and paid to complete the labeling through our system. The spamicity score has been calculated for each review and a label has been assigned to every review based on their spamicity score. The Fleiss's Kappa is calculated for three annotators with value of 0.89, which shows "almost perfect agreement" between them.

The labeled real-life product review dataset is the second contribution of this study. To test the accuracy of our model, we also re-labeled a portion of available Yelp.com dataset through our system and calculated the disagreement with their actual label based on the Yelp.com's filtering system. We found that only 7% of the reviews were labeled differently.

The other open problem of fake product review classification is the lack of historic knowledge independent feature sets. Most of the feature-based fake review detection techniques are only applicable on a specific product domain or historic knowledge is needed to extract these features. To address the problem, this study presents a set of domain and historic knowledge independent features, namely writing style and readability, which can be applied to almost any review hosting site. The feature set is the third contribution of this study. Writing style here refers to linguistic aspects that identify fake and real reviewers. Fake reviewers try hard to write a review that sounds like genuine, hence it affects their writing style and also readability of their fake reviews consequently. The method dependently detects reviewers' writing style before spamming can hurt a product or a business. The evaluation results of our features on the only available crowdsourced labeled gold standard dataset, with the accuracy of 90.7%, and on our proposed dataset with the accuracy of 98.9%, suggest significant differences between fake and real reviews on writing style and readability level.

Abstrak tesis yang dikemukakan kepada Senat Universiti Putra Malaysia sebagai memenuhi keperluan untuk ijazah Doktor Falsafah

MODEL ANOTASI ULASAN PALSU DAN KLASIFIKASI MELALUI GAYA PENULISAN PENGULAS

Oleh

SOMAYEH SHOJAEE

September 2019

Pengerusi: Profesor Madya Masrah Azrifah Azmi Murad, PhD
Fakulti: Sains Komputer dan Teknologi Maklumat

Dalam dekad yang lalu, ulasan produk dalam talian telah menjadi sumber utama maklumat semasa membuat keputusan dan proses pembelian perniagaan. Malangnya, penipu telah menghasilkan ulasan yang tidak benar secara sengaja untuk keuntungan atau publisiti. Aktiviti mereka menipu organisasi berpotensi untuk membentuk semula perniagaan mereka, pelanggan daripada membuat keputusan terbaik, dan teknik perlombongan pendapat dari mencapai kesimpulan yang tepat.

Salah satu cabaran besar untuk mengkaji ulasan spam ialah kekurangan set data ulasan produk sebenar yang dilabel. Kajian pelabelan produk secara manual sebagai palsu atau nyata adalah salah satu pendekatan untuk menangani masalah ini. Walau bagaimanapun, mengenali sama ada ulasan adalah palsu atau sebenar adalah sukar dengan hanya membaca kandungan ulasan, kerana *spammer* dengan mudah boleh membuat ulasan palsu yang sama seperti ulasan sebenar yang lain.

Untuk menangani masalah ini, perjanjian antara penganotasi dalam pendekatan pelabelan secara manual dipertingkat dengan membangunkan satu model baharu untuk menganotasi ulasan produk sebagai palsu atau sebenar. Inilah sumbangan pertama kajian penyelidikan ini. Model yang dicadangkan direka, dilaksanak dan diakses secara dalam talian. Ulasan kami yang didapati melalui merangkak dilabelkan oleh tiga penganotasi yang dilatih dan dibayar untuk melengkapkan pelabelan melalui sistem kami. Skor kerumitan telah dikira untuk setiap ulasan dan setiap label telah di berikan kepada

setiap ulasan berdasarkan skor *spamicity* mereka. Fleiss's Kappa dihitung untuk tiga penganotasi dengan nilai 0.89, yang menunjukkan perjanjian yang hampir sempurna di antara mereka. Set data ulasan produk sebenar yang dilabelkan adalah sumbangan kedua kajian ini. Untuk menguji ketepatan model kami, kami juga melabel semula sebahagian daripada set data Yelp.com yang tersedia melalui sistem kami dan dikira ketidaksepakatan dengan label sebenar mereka berdasarkan sistem pengenalan Yelp.com. Kami mendapati bahawa hanya 7% daripada ulasan dilabelkan secara berbeza.

Masalah terbuka yang lain untuk klasifikasi ulasan produk palsu adalah kekurangan set ciri bebas bersejarah. Kebanyakan teknik pengesanan ulasan palsu berasaskan ciri hanya boleh digunakan pada domain produk spesifik atau pengetahuan bersejarah diperlukan untuk mengekstrak ciri-ciri ini. Untuk menangani masalah ini, kajian ini membentangkan satu set ciri domain dan ciri bebas bersejarah, iaitu gaya penulisan dan kebolehbacaan yang boleh digunakan untuk mana-mana laman tapak pengehosan ulasan. Set ciri ini ialah sumbangan ketiga kajian ini. Gaya penulisan di sini merujuk kepada aspek linguistik yang mengenal pasti pengulas palsu dan sebenar. Pengulas palsu berusaha keras menulis ulasan berbunyi seperti tulen, oleh itu ia serupa dengan gaya tulisan mereka dan juga kebolehbacaan ulasan mereka yang palsu. Kaedah ini bergantung pada gaya penulisan pengulas sebelum spam yang boleh menjejaskan produk atau perniagaan. Keputusan penilaian ciri-ciri kami pada set data yang disediakan oleh orang ramai dengan ketepatan 90.7%, dan pada set data yang dicadangkan dengan ketepatan 98.9%, mencadangkan perbezaan antara ulasan palsu dan sebenar mengenai gaya penulisan dan tahap kebolehbacaan.

ACKNOWLEDGEMENTS

First and foremost, I would like to express my sincere gratitude to my thesis supervisor, *Assoc. Prof. Dr. Masrah Azrifah Azmi Murad*, for all her support and patience during the years of my PhD studies and for trusting me to pretty much direct my own research.

I would like to thank partial financial support from the Ministry of Science, *Technology and Innovation Malaysia (MOSTI)*, Project Number: 5450721, and *Research University Grant Scheme (RUGS)*, Universiti Putra Malaysia, Project Number: 05-02-12-2153RU.

I also would like to thank my committee members, *Assoc. Prof. Dr. Azreen bin Azman*, *Assoc. Prof. Dr. Nurfadhlina Mohd Sharef* for their valuable suggestions and advises through this work.

My very special thanks would certainly go to my loving and devoted *mother* for showing faith in me and her unconditional love, encouragement and support throughout my life. I would like to thank my father and mother in law for their support through my PhD too.

I am most grateful to my husband, *Saeid*, for his companionship on my journey. Our true love is the main source of strength, motivation and inspiration. He always supports me when I thought that it is impossible to continue and I deeply appreciate his belief in me. Last but not least, I appreciate my precious little angel, *Hannah*, from heaven above for the patience she showed during my study. Her smiles always gave me strength to complete my PhD. Words would never say how grateful I am to both of them. I am the world luckiest person to have such a caring and lovely family, who always stand beside me with their unconditional love and support.

This thesis was submitted to the Senate of Universiti Putra Malaysia and has been accepted as fulfilment of the requirement for the degree of Doctor of Philosophy.

The members of the Supervisory Committee were as follows:

Masrah Azrifah Azmi Murad, PhD
Associate Professor
Faculty of Computer Science and Information Technology
Universiti Putra Malaysia
(Chairperson)

Azreen Bin Azman, PhD
Associate Professor
Faculty of Computer Science and Information Technology
Universiti Putra Malaysia
(Member)

Nurfadhlina Mohd Sharef, PhD
Associate Professor
Faculty of Computer Science and Information Technology
Universiti Putra Malaysia
(Member)

ZALILAH BINTI SHARIFF, PhD
Professor and Deputy Dean
School of Graduate Studies
Universiti Putra Malaysia

Date:

Declaration by Graduate Student

I hereby confirm that:

- this thesis is my original work;
- quotations, illustrations and citations have been duly referenced;
- this thesis has not been submitted previously or concurrently for any other degree at any other institutions;
- intellectual property from the thesis and copyright of thesis are fully-owned by Universiti Putra Malaysia, as according to the Universiti Putra Malaysia (Research) Rules 2012;
- written permission must be obtained from supervisor and the office of Deputy Vice-Chancellor (Research and Innovation) before thesis is published (in the form of written, printed or in electronic form) including books, journals, modules, proceedings, popular writings, seminar papers, manuscripts, posters, reports, lecture notes, learning modules or any other materials as stated in the Universiti Putra Malaysia (Research) Rules 2012;
- there is no plagiarism or data falsification/fabrication in the thesis, and scholarly integrity is upheld as according to the Universiti Putra Malaysia (Graduate Studies) Rules 2003 (Revision 2012-2013) and the Universiti Putra Malaysia (Research) Rules 2012. The thesis has undergone plagiarism detection software.

Signature: ______________ Date: 25 January 2020

Name and Matric No.: Somayeh Shojaee, GS33458

Declaration by Members of Supervisory Committee

This is to confirm that:

- the research conducted and the writing of this thesis was under our supervision;
- supervision responsibilities as stated in the Universiti Putra Malaysia (Graduate Studies) Rules 2003 (Revision 2012-2013) are adhered to.

Signature: ______________________
Name of Chairman
of Supervisory
Committee: Masrah Azrifah Azmi Murad

Signature: ______________________
Name of Member
of Supervisory
Committee: Azreen Bin Azman

Signature: ______________________
Name of Member
of Supervisory
Committee: Nurfadhlina Mohd Sharef

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AMT	Amazon Mechanical Turk
ARI	Automated Readability Index
AUC	Area Under Curve
CLI	Coleman-Liau Index
DM	Data Mining
FN	False Negative
FP	False Positive
FPRD	Fake Product Review Detection
GSRank	Group Spam Rank
ICF	Iterative Computation Framework
ILM	Inferential Language Modelling
IR	Information Retrieval
JSD	Jensen-Shannon Divergence
KLD	Kullback-Leibler Divergence
LCM	Latent Collusion Model
LDA	Latent Dirichlet Allocation
LIWC	Linguistic Inquiry and Word Count
LR	Logistic Regression
LSM	Latent Spam Model
MAE	Mean Absolute Error
MAUE	Mean Absolute User Error
MHCC	Multi-typed Heterogeneous Collective Classification
NB	Naive Bayes
NDCG	Normalized Discounted Cumulative Gain
NLP	Natural Language Processing
NLTK	Natural Language Toolkit
OM	Opinion Mining
PCFG	Probabilistic Context-Free Grammars
POS	Part Of Speech
PU-learning	Positive Unlabeled-learning
RWR	Random Walk with Restarts
SL	Simple Logistic
SMO	Sequential Minimal Optimization
SMOG	Simple Measure of Gobbledygook
SVM	Support Vector Machine
SWM	Software Marketplace
TF-IDF	Term Frequency-Inverse Document Frequency
TN	True Negative
TP	True Positive
UCB	Upper Confident Bound

CHAPTER 1

INTRODUCTION

1.1 Motivation

Online reviews are the main source of information during customers and businesses decision making processes. However, abundance of fake reviews, are written by someone who has not actually used the product or service, causes loss of trust on product reviews and also prevents Opinion Mining (OM) techniques from reaching accurate conclusions.

Fake review detection is an interrelated area of research which uses various techniques taken from Natural Language Processing (NLP), Information Retrieval (IR), and Data Mining (DM) (Ravi and Ravi, 2015). Detecting fake review is a multi-faceted problem and different steps are needed to perform fake review detection from a given set of reviews, since reviews come from several sources and in various formats. Data acquisition, data preprocessing, data labeling, feature generation and spamming detection techniques are the most regular sub-tasks that are required to be performed to detect fake reviews.

Data collection and labeling are the main part of any machine learning based model. Collecting huge amount of reviews and labeling an adequate number of them to train a classifier is a challenging task. Many researches have been done since 2007 when the first article was published in the area of fake product review detection by Jindal and Liu (2007b) and it is still a hot topic.

One of the biggest challenges in Fake Product Review Detection (FPRD) is lack of publicly available gold standard real-life labeled product review dataset. One of the approaches to annotate review datasets is labeling manually a portion of real review corpus by hiring annotators. Manually labeling product reviews as spam and non-spam was presented by Jindal and Liu (2008) for the first time. Identifying spam reviews among lots of reviews is extremely hard by only reading the review because spammer can craft a spam opinion which is the same as any other non-spam review.

Second approach to create a labeled fake review dataset is hiring people to craft fake reviews through crowdsourcing websites or paying professionals to write about the product, for instance a hotel owner can ask the hotel's employees to write fake reviews about his hotel. The problem with this approach is gathering a big amount of artificial reviews is almost impossible and they are not the same as the real-life reviews (Mukherjee et al., 2013b), hence the results of studies based on these data may differ from real-life data.

Recently, some studies, e.g. Rayana and Akoglu (2015), collect spam reviews from "Yelp.com", which has a fake review filtering system. The problem with these data, namely commercial labeled data, is no publicly available information about the Yelp.com's filtering system. Some other studies , e.g. Banerjee et al. (2015), collect genuine reviews from "Hotels.com", "Expedia.com" and "Agoda.com", which only allow people to leave a comment only if they paid and stayed in the hotel. Collecting spam reviews still is there when using the three mentioned review hosting websites. Only genuine reviews can be collected from these websites and there is no access to the filtered reviews. Commercial labeled data also has review hosting site dependency and it is not a method to label any reviews from any website.

Liu (2015b) categorized three main types of data for FPRD. By the help of these data, many spam features are engineered to spot fake reviews. The data types are:

1. **Review content:** which includes text content of each review.
2. **Meta-data about the review:** which are collected and accessible only by businesses' websites not to researchers. Some example of these data are, geographic location of the reviewer and sequence of reviewers' clicking. This kind of information is very suitable for reviewers' abnormal behavioural patterns recognition.
3. **Product information:** which can be collected through e.g. the product description or sales rank/volume. For instance, reviews of a product with lots of positive reviews and few numbers of sell are more likely fake.

Although machine learning techniques are not 100% efficient on FPRD, they are more accurate than manual detection (Crawford et al., 2016). As mentioned by Abbasi et al. (2010), distinguishing words (features) can give ultimate indication for spam and non-spam review classification. Feature based machine learning methods are pretty effective to spot fake reviews and reviewers. One of the common techniques in text mining is using bag-of-words, where a bunch of individual words, or small sets of words, are used as features; however, regarding several research results this method is not adequate to train a classifier with sufficient efficiency in FPRD (Abbasi et al., 2010). In the literature, there are many research studies that discuss different feature sets for the purpose of FPRD by applying a variety of machine learning techniques such as classification. For example, Jindal et al. (2010), Li et al.

(2011) and Mukherjee et al. (2012) used words from reviews' content as the features. Another research by Ott et al. (2013) applied review characteristic features, as well as unigram and bigram term-frequencies.

In this thesis, we discuss machine learning techniques that are applied to FPRD by emphasizing on feature extraction and those features' impacts on the performance of the fake review detectors. Most of the existing supervised and semi-supervised learning techniques on review spam detection focus on domain dependent features. Using those features such as "total helpful feedback number" (Jindal and Liu, 2008), or features like Part Of Speech (POS) are the most common techniques that are presented in previous studies. These kinds of features are applicable for specific review hosting sites or data. For instance, number of mentioned specific words such as "clean" using POS on hotel reviews can't help on other products such as DVD. Hence, there is the lack of research on domain independent techniques that can spot fake reviews.

1.2 Problem Statements

We define the FPRD problem as a classification task, to spot fake and real reviews. The accuracy of classification depends heavily on the discriminatory power of the extracted features and classifiers' evaluation on real world labeled data. In spite of significant advances in FPRD, there are still some challenging open problems. In this thesis, the problems of FPRD that need to be tackled are investigated separately in two research issues as follows:

1. Manually labeling a portion of product review corpus by hiring annotators is one of the common approaches to prepare a labeled dataset. This approach is applied in many studies such as Xue et al. (2015); Lu et al. (2013); Huang et al. (2013); Xie et al. (2012a); Wang et al. (2012); Feng et al. (2012b); Mukherjee et al. (2012); Lau et al. (2011); Lai et al. (2010); Jindal et al. (2010) and Jindal and Liu (2008, 2007b,a). However, recognizing whether a review is fake from real-life data is very difficult by just reading the review content because a spammer can carefully craft a fake review which is like any other real review. Therefore, the inter-annotator agreement is usually low.
2. One research direction of FPRD is to explore feature sets. Particularly, textual features that can overcome the limitations of historic knowledge dependent features (behavioural, profile and product based features) (Jindal and Liu, 2007a,b, 2008; Mukherjee et al., 2013a; Zhang et al., 2014; Rayana and Akoglu, 2015; Li et al., 2016b; Shehnepoor et al., 2017). However, although historic knowledge dependent features are

useful in FPRD, their historic knowledge dependency as well as public unavailability of meta-data for most of the review hosting websites are among the limitations that restrict their usage in FPRD.

1.3 Research Objectives

The goal of this study is to propose a model and new features to cover relevant aspects and issues of FPRD by improving data annotation and classification problems. The research objectives of this thesis can hence be summarized as follows:

1. To propose, implement and evaluate a fake product review annotation model to improve the inter-annotator agreement in manually labeling technique.
2. To identify a historic knowledge independent feature set that improves the accuracy of fake product review classification.

The Figure 1.1 shows the correlation between the objectives. After crawling the review corpus and labeling them via the annotating system, we applied the labeled corpus on our historic independent feature set. Finally we have prepared a dataset which includes historic knowledge independent real-life product review.

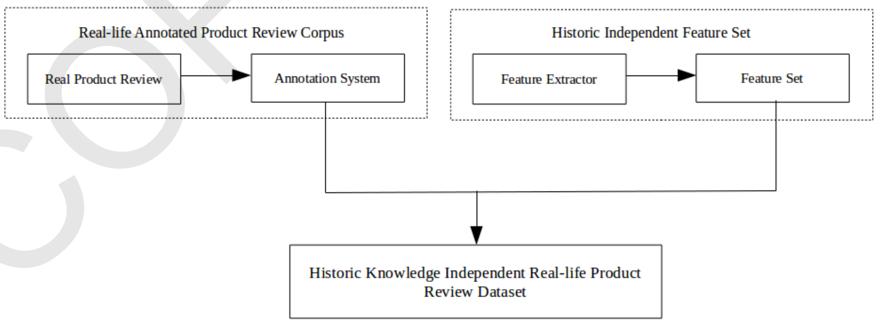

Figure 1.1: Correlation Between the Objectives

1.4 Research Scope

The scope of this research is limited to fake product review collection, annotation and supervised feature-based spam review detection techniques. In this study, we only choose spam product reviews and among all three types of labeled data collection approaches, namely manually, crowdsourcing and commercial, we focused and improved the manually labeling approach. To detect fake reviews, we proposed linguistic-based feature sets to apply the supervised classification techniques.

1.5 Research Contributions

The contributions of this thesis are innovative approaches that address the above-mentioned research issues as follows:

- A model to annotate product reviews as fake or real to enhance the inter-annotator agreement in manually labeling technique.
- A new labeled product review dataset to improve the evaluation process of fake product review classification.
- A new historic knowledge independent feature set, namely the authors' writing style and reviews' readability features, to improve fake product review classification.

1.6 Organization of Thesis

This thesis is organized as follows. Chapter 2 provides the literature review of related works in FPRD area. Existing datasets, features and detection approaches are presented in this chapter. Research methodology is discussed in Chapter 3, including research overview, experimental design and the steps of research study. The contributions of this study are presented in Chapters 4 and 5. Chapter 4 describes the proposed annotation model, as well as the implementation and evaluation of the model and our new dataset. Chapter 5 describes the details of writing style and readability features, feature extraction process, classification evaluation and comparison with the baseline study. Finally, Chapter 6 concludes the thesis and discusses future work.

REFERENCES

Abbasi, A., Zhang, Z., Zimbra, D., Chen, H. and Nunamaker, J. F. 2010. Detecting Fake Websites: The Contribution of Statistical Learning Theory. *MIS Quarterly* 34 (3): 435–461.

Afroz, S., Brennan, M. and Greenstadt, R. 2012. Detecting Hoaxes, Frauds, and Deception in Writing Style Online, 461–475.

Ahsan, M. N. I., Nahian, T., Kafi, A. A., Hossain, M. I. and Shah, F. M. 2016a. An ensemble approach to detect review spam using hybrid machine learning technique. In *2016 19th International Conference on Computer and Information Technology (ICCIT)*, 388–394.

Ahsan, M. N. I., Nahian, T., Kafi, A. A., Hossain, M. I. and Shah, F. M. 2016b. Review spam detection using active learning. In *2016 IEEE 7th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON)*, 1–7.

Akoglu, L., Chandy, R. and Faloutsos, C. 2013. Opinion Fraud Detection in Online Reviews by Network Effects. In *ICWSM*.

Algur, S. P., Patil, A. P., Hiremath, P. S. and Shivashankar, S. 2010. Conceptual Level Similarity Measure Based Review Spam Detection. In *2010 International Conference on Signal and Image Processing (ICSIP)*, 416–423.

Amigó, E., Gonzalo, J., Artiles, J. and Verdejo, F. 2009. A Comparison of Extrinsic Clustering Evaluation Metrics Based on Formal Constraints. *Information Retrieval* 12 (4): 461–486.

Baayen, H., van Halteren, H. and Tweedie, F. 1996. Outside the Cave of Shadows: Using Syntactic Annotation to Enhance Authorship Attribution. *Literary and Linguistic Computing* 11 (3): 121–132.

Banerjee, S. and Chua, A. Y. 2017. Authentic versus fictitious online reviews: A textual analysis across luxury, budget, and mid-range hotels. *Journal of Information Science* 43 (1): 122–134.

Banerjee, S. and Chua, A. Y. K. 2014a. A Linguistic Framework to Distinguish between Genuine and Deceptive Online Reviews.

Banerjee, S. and Chua, A. Y. K. 2014b. A Theoretical Framework to Identify Authentic Online Reviews. *Online Information Review* 38 (5): 634–649.

Banerjee, S. and Chua, A. Y. K. 2014c. Applauses in hotel reviews: Genuine or deceptive? In *2014 Science and Information Conference*, 938–942.

Banerjee, S., Chua, A. Y. K. and Kim, J.-j. 2015. Let's vote to classify authentic and manipulative online reviews: The role of comprehensibility, informativeness and writing style. In *Science and Information Conference (SAI), 2015*, 77–83.

Banerjee, S., Chua, A. Y. K. and Kim, J.-J. 2017. Don't be deceived: Using linguistic analysis to learn how to discern online review authenticity. *Journal of the Association for Information Science and Technology* 68 (6): 1525–1538.

Banerjee, S., Chua, A. Y. K. and Kim, W. 2014. Dissecting Genuine and Deceptive Kudos: The Case of Online Hotel Reviews. In *International Journal of Advanced Computer Science and Applications(IJACSA), Special Issue on Extended Papers from Science and Information Conference 2014*.

Blei, D. M., Ng, A. Y. and Jordan, M. I. 2003. Latent Dirichlet Allocation. *J. Mach. Learn. Res.* 3: 993–1022.

Blitzer, J., Dredze, M. and Pereira, F. 2007. Biographies, Bollywood, Boomboxes and Blenders: Domain Adaptation for Sentiment Classification. In *In ACL*, 187–205.

Buhrmester, M., Kwang, T. and Gosling, S. D. 2011. Amazon's Mechanical Turk: A New Source of Inexpensive, Yet High-Quality, Data? *Perspectives on Psychological Science* 6 (1): 3–5. PMID: 26162106.

Burgoon, J. K., Blair, J. P., Qin, T. and Nunamaker, J. F. 2003. *Detecting Deception through Linguistic Analysis*, 91–101. Berlin, Heidelberg: Springer Berlin Heidelberg.

Cagnina, L. C. and Rosso, P. 2015. Classification of deceptive opinions using a low dimensionality representation. In *Proceedings of the 6th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis*, 58–66.

Chang, T., Hsu, P. Y., Cheng, M. S., Chung, C. Y. and Chung, Y. L. 2015. Detecting Fake Review with Rumor Model–Case Study in Hotel Review. In *Revised Selected Papers, Part II, of the 5th International Conference on Intelligence Science and Big Data Engineering. Big Data and Machine Learning Techniques - Volume 9243*, 181–192. New York, NY, USA: Springer-Verlag New York, Inc.

Chen, C., Zhao, H. and Yang, Y. 2015. Deceptive Opinion Spam Detection Using Deep Level Linguistic Features. In *Natural Language Processing and Chinese Computing* (eds. J. Li, H. Ji, D. Zhao, and Y. Feng), 465–474. Springer International Publishing.

Chen, F., Tan, P.-N. and Jain, A. K. 2009. A Co-classification Framework for Detecting Web Spam and Spammers in Social Media Web Sites, 1807–1810. New York, NY, USA: ACM.

Chengzhang, J. and Kang, D.-k. 2015. Detecting the spam review using tri-training. In *2015 17th International Conference on Advanced Communication Technology (ICACT)*, 374–377.

Chino, D. Y. T., Costa, A. F., Traina, A. J. M. and Faloutsos, C. 2017. *VOLTIME: Unsupervised Anomaly Detection on Users Online Activity Volume*, 108–116. Society for Industrial and Applied Mathematics.

Cocarascu, O. and Toni, F. 2016. Detecting Deceptive Reviews Using Argumentation. In *Proceedings of the 1st International Workshop on AI for Privacy and Security*, 9:1–9:8. New York, NY, USA: ACM.

Coleman, M. and Liau, T. L. 1975. A Computer Readability Formula Designed for Machine Scoring. *Journal of Applied Psychology* 60 (2): 283.

Crawford, M., Khoshgoftaar, T. and Prusa, J. 2016. Reducing Feature Set Explosion to Facilitate Real-World Review Spam Detection. In *Florida Artificial Intelligence Research Society Conference*.

Dale, E. and Chall, J. S. 1948. A Formula for Predicting Readability: Instructions. *Educational Research Bulletin* 27 (2): 37–54.

Dang, T. N. Y. and Webb, S. 2016. *Making an essential word list for beginners*, Ch. Chapter 15: Making an essential word list for beginners, 153–167. John Benjamins.

Das, S. R., Chen, M. Y., Agarwal, T. V., Brooks, C., shee Chan, Y., Gibson, D., Leinweber, D., Martinez-jerez, A., Raghubir, P., Rajagopalan, S., Ranade, A., Rubinstein, M. and Tufano, P. 2001. Yahoo! for Amazon: Sentiment Extraction from Small Talk on the Web. In *8th Asia Pacific Finance Association Annual Conference*.

Dave, K., Lawrence, S. and Pennock, D. M. 2003. Mining the Peanut Gallery: Opinion Extraction and Semantic Classification of Product Reviews. In *Proceedings of the 12th International Conference on World Wide Web*, 519–528. New York, NY, USA: ACM.

Dewang, R. K., Singh, P. and Singh, A. K. 2016. Finding of Review Spam Through "Corleone, Review Genre, Writing Style and Review Text Detail Features". In *Proceedings of the Second International Conference on Information and Communication Technology for Competitive Strategies*, 23:1–23:6. New York, NY, USA: ACM.

Ding, W., Shang, Y., Park, D. H., Guo, L. and Hu, X. 2015. *ProbitUCB: A Novel Method for Review Ranking*, 3–15. Cham: Springer International Publishing.

Eisenstein, J., Ahmed, A. and Xing, E. P. 2011. Sparse Additive Generative Models of Text. In *Proceedings of the 28th International Conference on International Conference on Machine Learning*, 1041–1048. USA: Omnipress.

Fayazi, A., Lee, K., Caverlee, J. and Squicciarini, A. 2015. Uncovering Crowdsourced Manipulation of Online Reviews. In *Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval*, 233–242. New York, NY, USA: ACM.

Fearis, B. 2007, Authenticity of Hotel Reviews Questioned.

Fei, G., Mukherjee, A., Liu, B., Hsu, M., Castellanos, M. and Ghosh, R. 2013. Exploiting Burstiness in Reviews for Review Spammer Detection.

Feng, S., Banerjee, R. and Choi, Y. 2012a. Syntactic Stylometry for Deception Detection, 171–175. Stroudsburg, PA, USA: Association for Computational Linguistics.

Feng, S., Xing, L., Gogar, A. and Choi, Y. 2012b. Distributional Footprints of Deceptive Product Reviews. In *Sixth International AAAI Conference on Weblogs and Social MediaI (CWSM)*, 98–105.

Feng, V. W. and Hirst, G. 2013. Detecting Deceptive Opinions with Profile Compatibility. In *IJCNLP*.

Ferraz Costa, A., Yamaguchi, Y., Juci Machado Traina, A., Traina, Jr., C. and Faloutsos, C. 2015. RSC: Mining and Modeling Temporal Activity in Social Media. In *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 269–278. New York, NY, USA: ACM.

Fusilier, D. H., Cabrera, R. G., Montes-y Gomez, M. and Rosso, P. 2013. Using PU-Learning to Detect Deceptive Opinion Spam. In *The 4th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis*, 38–45.

Fusilier, D. H., y Gmez, M. M., Rosso, P. and Cabrera, R. G. 2015a. Detecting positive and negative deceptive opinions using PU-learning. *Information Processing & Management* 51 (4): 433 – 443.

Fusilier, D. H., y Gmez, M. M., Rosso, P. and Cabrera, R. G. 2015b. Detection of Opinion Spam with Character n-grams. In *Computational Linguistics and Intelligent Text Processing* (ed. A. Gelbukh), 285–294. Cham: Springer International Publishing.

Goswami, K., Park, Y. and Song, C. 2017. Impact of reviewer social interaction on online consumer review fraud detection. *Journal of Big Data* 4 (1): 15.

Gunning, R. 1969. The Fog Index After Twenty Years. *Journal of Business Communication* 6 (2): 3–13.

Hand, D. J., Smyth, P. and Mannila, H. 2001. *Principles of Data Mining*. Cambridge, MA, USA: MIT Press.

Harris, C. G. 2012. Detecting Deceptive Opinion Spam Using Human Computation. In *Workshops at the Twenty-Sixth AAAI Conference on Artificial Intelligence*, 87–93. AAAI Press.

He, R. and McAuley, J. 2016. Ups and Downs: Modeling the Visual Evolution of Fashion Trends with One-Class Collaborative Filtering. In *Proceedings of the 25th International Conference on World Wide Web*, 507–517. Republic and Canton of Geneva, Switzerland: International World Wide Web Conferences Steering Committee.

Heredia, B., Khoshgoftaar, T. M., Prusa, J. D. and Crawford, M. 2017. Improving detection of untrustworthy online reviews using ensemble learners combined with feature selection. *Social Network Analysis and Mining* 7 (1): 37.

Heydari, A., Tavakoli, M. and Salim, N. 2016. Detection of fake opinions using time series. *Expert Systems with Applications* 58: 83 – 92.

Hoang, N. 2014. Opinion Spam Recognition Method for Online Reviews Using Ontological Features. *Tap chi KHOA HoC DHSP TPHCM* 61: 44–59.

Hooi, B., Shah, N., Beutel, A., Gunneman, S., Akoglu, L., Kumar, M., Makhija, D. and Faloutsos, C. 2015, BIRDNEST: Bayesian Inference for Ratings-Fraud Detection, arXiv:1511.06030 [cs.AI].

Hosseinia, M. and Mukherjee, A. 2017. Detecting Sockpuppets in Deceptive Opinion Spam. In *proceedings of the 18th International Conference on Intelligent Text Processing and Computational Linguistics (CICLING'17)*.

Hotchkiss, S. N. and Paterson, D. G. 1951. *How to Test Readability.*. Harper Row.

Hu, M. and Liu, B. 2004. Mining and Summarizing Customer Reviews. In *Proceedings of the Tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 168–177. New York, NY, USA: ACM.

Huang, C., Jiang, Q. and Zhang, Y. 2010. Detecting Comment Spam through Content Analysis. In *Web-Age Information Management* (eds. H. Shen, J. Pei, M. T. Özsu, J. Zou Leiand Lu, T.-W. Ling, G. Yu, Y. Zhuang, and J. Shao), 222–233. Jiuzhaigou Valley, China: Springer Berlin Heidelberg.

Huang, J., Qian, T., He, G., Zhong, M. and Peng, Q. 2013. *Detecting Professional Spam Reviewers*, 288–299. Berlin, Heidelberg: Springer Berlin Heidelberg.

Iqbal, F., Binsalleeh, H., Fung, B. C. M. and Debbabi, M. 2010. Mining Writeprints From Anonymous e-mails for Forensic Investigation. *Digit. Investig.* 7 (1-2): 56–64.

Järvelin, K. and Kekäläinen, J. 2000. IR Evaluation Methods for Retrieving Highly Relevant Documents. In *Proceedings of the 23rd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, 41–48. New York, NY, USA: ACM.

Jindal, N. and Liu, B. 2007a. Analyzing and Detecting Review Spam. In *Seventh IEEE International Conference on Data Mining (ICDM 2007)*, 547–552. Omaha, NE, USA: Ieee.

Jindal, N. and Liu, B. 2007b. Review Spam Detection. In *Proceedings of the 16th International Conference on World Wide Web*, 1189–1190. New York, NY, USA: ACM.

Jindal, N. and Liu, B. 2008. Opinion Spam and Analysis, 219–230. New York, NY, USA: ACM.

Jindal, N., Liu, B. and Lim, E.-P. 2010. Finding Unusual Review Patterns Using Unexpected Rules, 1549–1552. New York, NY, USA: ACM.

Jo, Y. and Oh, A. H. 2011. Aspect and Sentiment Unification Model for Online Review Analysis. In *Proceedings of the fourth ACM international conference on Web search and data mining*, 815–824. New York, NY, USA: ACM.

Junting Ye, Santhosh Kumar, L. A. 2016. Temporal Opinion Spam Detection by Multivariate Indicative Signals. In *Proceedings of the Tenth International AAAI Conference on Web and Social Media (ICWSM 2016)*.

Kawate, S. and Patil, K. 2017. An Approach for Reviewing and Ranking the Customers' Reviews through Quality of Review (QoR). *ICTACT Journal on Soft Computing* .

KC, S. and Mukherjee, A. 2016. On the Temporal Dynamics of Opinion Spamming: Case Studies on Yelp. In *Proceedings of the 25th International Conference on World Wide Web*, 369–379. Republic and Canton of Geneva, Switzerland: International World Wide Web Conferences Steering Committee.

Kim, S., Chang, H., Lee, S., Yu, M. and Kang, J. 2015. Deep Semantic Frame-Based Deceptive Opinion Spam Analysis. In *Proceedings of the 24th ACM International on Conference on Information and Knowledge Management*, 1131–1140. New York, NY, USA: ACM.

Kim, S., Lee, S., Park, D. and Kang, J. 2017. Constructing and Evaluating a Novel Crowdsourcing-based Paraphrased Opinion Spam Dataset. In *Proceedings of the 26th International Conference on World Wide Web*, 827–836. Republic and Canton of Geneva, Switzerland: International World Wide Web Conferences Steering Committee.

Kincaid, J. P. and Delionbach, L. J. 1973. Validation of the Automated Readability Index: A Follow-Up. *Human Factors* 15 (1): 17–20.

Kincaid, J. P., Jr., R. P. F., Rogers, R. L. and Chissom, B. S. 1975, Derivation of New Readability Formulas (Automated Readability Index, Fog Count, and Flesch Reading Ease Formula) for Navy Enlisted Personnel, Tech. rep., Chief of Naval Technical Training: Naval Air Station Memphis.

Kullback, S. and Leibler, R. A. 1951. On Information and Sufficiency. *The Annals of Mathematical Statistics* 22 (1): 79–86.

Lafferty, J. and Zhai, C. 2001. Document Language Models, Query Models, and Risk Minimization for Information Retrieval. In *Proceedings of the 24th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, 111–119. New York, NY, USA: ACM.

Lai, C. L., Xu, K. Q., Lau, R. Y. K., Li, Y. and Song, D. 2010. High-Order Concept Associations Mining and Inferential Language Modeling for Online Review Spam Detection, 1120–1127.

Landis, J. R. and Koch, G. G. 1977. The Measurement of Observer Agreement for Categorical Data. *Biometrics* 33 (1): 159–174.

Lappas, T. 2012. *Fake Reviews: TheMalicious Perspective*, *Lecture Notes in Computer Science*, vol. 7337, Ch. Fake Revie, 23–34. Springer Berlin Heidelberg.

Lappas, T., Sabnis, G. and Valkanas, G. 2016. The Impact of Fake Reviews on Online Visibility: A Vulnerability Assessment of the Hotel Industry. *Information Systems Research* 27 (4): 940–961.

Lau, R. Y. K., Liao, S. Y., Kwok, R. C.-W., Xu, K., Xia, Y. and Li, Y. 2011. Text Mining and Probabilistic Language Modeling for Online Review Spam Detection. *ACM Trans. Manage. Inf. Syst.* 2 (4): 21–25.

Laughlin, G. H. M. 1969. SMOG Grading-a New Readability Formula. *Journal of Reading* 12 (8): 639–646.

Layton, J. R. 1980. A Chart for Computing the Dale-Chall Readability Formula Above Fourth Grade Level. *Journal of Reading* 24 (3): 239–244.

Leedy, P. D. and Ormrod, J. E. 2015. *Practical Research: Planning and Design (11th Edition)*. Pearson.

Li, F., Huang, M., Yang, Y. and Zhu, X. 2011. Learning to Identify Review Spam, 2488–2493. Barcelona, Catalonia, Spain: AAAI Press.

Li, H., Chen, Z., Liu, B., Wei, X. and Shao, J. 2014a. Spotting Fake Reviews via Collective Positive-Unlabeled Learning. In *Proceedings of the 2014 IEEE International Conference on Data Mining*, 899–904. Washington, DC, USA: IEEE Computer Society.

Li, H., Chen, Z., Mukherjee, A., Liu, B. and Shao, J. 2015a. Analyzing and Detecting Opinion Spam on a Large-scale Dataset via Temporal and Spatial Patterns. In *Proceedings of Ninth International AAAI Conference on Web and Social Media*, 634–637.

Li, H., Fei, G., Wang, S., Liu, B., Shao, W., Mukherjee, A. and Shao, J. 2016a. Modeling Review Spam Using Temporal Patterns and Co-bursting Behaviors. *ArXiv e-prints* .

Li, H., Fei, G., Wang, S., Liu, B., Shao, W., Mukherjee, A. and Shao, J. 2017a. Bimodal Distribution and Co-Bursting in Review Spam Detection. In *Proceedings of the 26th International Conference on World Wide Web*, 1063–1072. Republic and Canton of Geneva, Switzerland: International World Wide Web Conferences Steering Committee.

Li, H., Liu, B., Mukherjee, A. and Shao, J. 2014b. Spotting Fake Reviews using Positive-Unlabeled Learning. *Computación y Sistemas* 18.

Li, J., Cardie, C. and Li, S. 2013a. TopicSpam: a Topic-Model based approach for spam detection. In *ACL*.

Li, J., Ott, M. and Cardie, C. 2013b. Identifying Manipulated Offerings on Review Portals. In *EMNLP*.

Li, J., Ott, M., Cardie, C. and Hovy, E. 2014c. Towards a General Rule for Identifying Deceptive Opinion Spam. In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, 1566–1576. Baltimore, Maryland: Association for Computational Linguistics.

Li, L., Qin, B., Ren, W. and Liu, T. 2017b. Document representation and feature combination for deceptive spam review detection. *Neurocomputing* 254: 33 – 41. Recent Advances in Semantic Computing and Personalization.

Li, L., Ren, W., Qin, B. and Liu, T. 2015b. *Learning Document Representation forDeceptive Opinion Spam Detection*, 393–404. Cham: Springer International Publishing.

Li, X. and Yan, X. 2016. *A Novel Chinese Text Mining Method for E-Commerce Review Spam Detection*, 95–106. Cham: Springer International Publishing.

Li, Y. 2015. Highlighting the Fake Reviews in Review Sequence with the Suspicious Contents and Behaviours. *Journal of Information and Computational Science* 12 (4): 1615–1627.

Li, Y., Feng, X. and Zhang, S. 2016b. Detecting Fake Reviews Utilizing Semantic and Emotion Model. In *2016 3rd International Conference on Information Science and Control Engineering (ICISCE)*, 317–320.

Lim, E.-P., Nguyen, V.-A., Jindal, N., Liu, B. and Lauw, H. W. 2010. Detecting Product Review Spammers Using Rating Behaviors, 939–948. New York, NY, USA: ACM.

Lin, Y., Zhu, T., Wu, H., Zhang, J., Wang, X. and Zhou, A. 2014. Towards online anti-opinion spam: Spotting fake reviews from the review sequence. In *2014 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM 2014)*, 261–264.

Liu, B. 2012. *Sentiment analysis and opinion mining*. USA: Morgan & Claypool Publishers.

Liu, B. 2015a. *Detecting Fake or Deceptive Opinions*, 259302. Cambridge University Press.

Liu, B. 2015b. *Sentiment analysis: Mining opinions, sentiments, and emotions*. Cambridge University Press.

Liu, B., Dai, Y., Li, X., Lee, W. S. and Yu, P. S. 2003. Building Text Classifiers Using Positive and Unlabeled Examples. In *Third IEEE International Conference on Data Mining*, 179–186.

Liu, J., Fu, X., Liu, L. and Yue, K. 2015. What Makes a Good Review: Analyzing Reviews on JD.com. In *2015 International Conference on Service Science (ICSS)*, 213–216.

Liu, X. and Croft, W. B. 2004. Cluster-based Retrieval Using Language Models. In *Proceedings of the 27th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, 186–193. New York, NY, USA: ACM.

Locke, L. F., Silverman, S. and Spirduso, W. W. 2010. *Reading and Understanding Research*. 3rd edn. SAGE PUBN.

Lu, Y., Zhang, L., Xiao, Y. and Li, Y. 2013. Simultaneously Detecting Fake Reviews and Review Spammers Using Factor Graph Model. In *Proceedings of the 5th Annual ACM Web Science Conference*, 225–233. New York, NY, USA: ACM.

Malbon, J. 2013. Taking Fake Online Consumer Reviews Seriously. *Journal of Consumer Policy* 36 (2): 139–157.

Manning, C. D., Raghavan, P. and Schütze, H. 2008. *Introduction to Information Retrieval*. New York, NY, USA: Cambridge University Press.

McAuley, J. and Leskovec, J. 2013. Hidden Factors and Hidden Topics: Understanding Rating Dimensions with Review Text. In *Proceedings of the 7th ACM Conference on Recommender Systems*, 165–172. New York, NY, USA: ACM.

McAuley, J., Targett, C., Shi, Q. and van den Hengel, A. 2015. Image-Based Recommendations on Styles and Substitutes. In *Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval*, 43–52. New York, NY, USA: ACM.

Minnich, A. J., Chavoshi, N., Mueen, A., Luan, S. and Faloutsos, M. 2015. TrueView: Harnessing the Power of Multiple Review Sites. In *Proceedings of the 24th International Conference on World Wide Web*, 787–797. New York, NY, USA: ACM.

Molla, A., Biadgie, Y. and Sohn, K.-A. 2018. *Detecting Negative Deceptive Opinion from Tweets*, 329–339. Singapore: Springer Singapore.

Morinaga, S., Yamanishi, K., Tateishi, K. and Fukushima, T. 2002. Mining Product Reputations on the Web. In *Proceedings of the Eighth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 341–349. New York, NY, USA: ACM.

Mosteller, F. and Wallace, D. L. 1964. *Inference and Disputed Authorship : The Federalist*. Reading, Mass. : Addison-Wesley.

Mukherjee, A., Liu, B. and Glance, N. 2012. Spotting Fake Reviewer Groups in Consumer Reviews. In *Proceedings of the 21st International Conference on World Wide Web*, 191–200. New York, NY, USA: ACM.

Mukherjee, A., Liu, B., Wang, J., Glance, N. and Jindal, N. 2011. Detecting group review spam, 93–94. New York, NY, USA: ACM.

Mukherjee, A. and Venkataraman, V. 2014, Opinion Spam Detection: An Unsupervised Approach using Generative Models, Tech. rep., UH.

Mukherjee, A., Venkataraman, V., Liu, B. and Glance, N. 2013a. Fake Review Detection: Classification and Analysis of Real and Pseudo Reviews.

Mukherjee, A., Venkataraman, V., Liu, B. and Glance, N. S. 2013b. What Yelp Fake Review Filter Might Be Doing?

Mukherjee, S., Dutta, S. and Weikum, G. 2016. *Credible Review Detection with Limited Information Using Consistency Features*, 195–213. Cham: Springer International Publishing.

Nasukawa, T. and Yi, J. 2003. Sentiment Analysis: Capturing Favorability Using Natural Language Processing. In *Proceedings of the 2Nd International Conference on Knowledge Capture*, 70–77. New York, NY, USA: ACM.

Ong, T., Mannino, M. and Gregg, D. 2014. Linguistic characteristics of shill reviews. *Electronic Commerce Research and Applications* 13 (2): 69–78.

Ott, M. 2011, Deceptive Opinion Spam Corpus v1.4.

Ott, M., Cardie, C. and Hancock, J. 2012. Estimating the Prevalence of Deception in Online Review Communities. In *Proceedings of the 21st international conference on World Wide Web - WWW '12*, 201. New York, New York, USA: ACM Press.

Ott, M., Cardie, C. and Hancock, J. T. 2013. Negative Deceptive Opinion Spam. Atlanta, Georgia, USA: Association for Computational Linguistics.

Ott, M., Choi, Y., Cardie, C. and Hancock, J. T. 2011. Finding Deceptive Opinion Spam by any Stretch of the Imagination, 309–319. Stroudsburg, PA, USA: Association for Computational Linguistics.

Pang, B., Lee, L. and Vaithyanathan, S. 2002. Thumbs Up?: Sentiment Classification Using Machine Learning Techniques. In *Proceedings of the ACL-02 Conference on Empirical Methods in Natural Language Processing - Volume 10*, 79–86. Stroudsburg, PA, USA: Association for Computational Linguistics.

Pearl, L. and Steyvers, M. 2012. Detecting Authorship Deception: A Supervised Machine Learning Approach Using Author Writeprints. *Literary and Linguistic Computing* 27 (2): 183–196.

Ponte, J. M. and Croft, W. B. 1998. A Language Modeling Approach to Information Retrieval. In *Proceedings of the 21st Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, 275–281. New York, NY, USA: ACM.

Qian, T. and Liu, B. 2013. *Identifying Multiple Userids of the Same Author*, 1124–1135. Association for Computational Linguistics (ACL).

Ravi, K. and Ravi, V. 2015. A Survey on Opinion Mining and Sentiment Analysis. *Know.-Based Syst.* 89 (C): 14–46.

Rayana, S. and Akoglu, L. 2015. Collective Opinion Spam Detection: Bridging Review Networks and Metadata. In *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 985–994. New York, NY, USA: ACM.

Rayana, S. and Akoglu, L. 2016. *Collective Opinion Spam Detection using Active Inference*, 630–638.

Reiter, C. 2007, Travel Web Sites Clamp Down on Bogus Reviews.

Ren, Y. and Ji, D. 2017. Neural networks for deceptive opinion spam detection: An empirical study. *Information Sciences* 385386: 213 – 224.

Ren, Y., Ji, D. and H., Z. 2014. Positive Unlabeled Learning for Deceptive Reviews Detection. In *Proceedings of First Conference on Empirical Methods in Natural Language Processing*, 488–498.

Ren, Y., Ji, D., Yin, L. and Zhang, H. 2015. Finding Deceptive Opinion Spam by Correcting the Mislabeled Instances. *Chinese Journal of Electronics* 24 (1): 52–57.

Rout, J. K., Singh, S., Jena, S. K. and Bakshi, S. 2016. Deceptive review detection using labeled and unlabeled data. *Multimedia Tools and Applications* 1–25.

Savage, D., Zhang, X., Yu, X., Chou, P. and Wang, Q. 2015. Detection of opinion spam based on anomalous rating deviation. *Expert Systems with Applications* 42 (22): 8650 – 8657.

Sharma, K. and Lin, K.-I. 2013. Review Spam Detector with Rating Consistency Check. In *Proceedings of the 51st ACM Southeast Conference*, 34:1–34:6. New York, NY, USA: ACM.

Shashirekha, H. L., Murali, S. and NAgabhushan, P. 2009. Ontology Based Similarity Measure for Text Documents. In *International conference on signal and image processing (ICSIP)*.

Shehnepoor, S., Salehi, M., Farahbakhsh, R. and Crespi, N. 2017. NetSpam: A Network-Based Spam Detection Framework for Reviews in Online Social Media. *IEEE Transactions on Information Forensics and Security* 12 (7): 1585–1595.

Solorio, T., Hasan, R. and Mizan, M. 2014. Sockpuppet Detection in Wikipedia: A Corpus of Real-World Deceptive Writing for Linking Identities. In *Proceedings of the Ninth International Conference on Language Resources and Evaluation, LREC 2014*, 1355–1358. Reykjavik, Iceland.

Stoppelman, J. 2009. *Why Yelp Has a Review Filter*.

Sun, C., Du, Q. and Tian, G. 2016. Exploiting Product Related Review Features for Fake Review Detection. *Mathematical Problems in Engineering* 2016 (1): 4935792.

Sun, H., Morales, A. and Yan, X. 2013. Synthetic Review Spamming and Defense. In *Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 1088–1096. New York, NY, USA: ACM.

Tan, P.-N. and Jin, R. 2004. Ordering Patterns by Combining Opinions from Multiple Sources. In *Proceedings of the Tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 695–700. New York, NY, USA: ACM.

Tong, R. 2001. An Operational System for Detecting and Tracking Opinions in on-line Discussions. In *Working Notes of the SIGIR Workshop on Operational Text Classification*, 1–6. New Orleans, Louisianna.

Turney, P. D. 2002. Thumbs Up or Thumbs Down?: Semantic Orientation Applied to Unsupervised Classification of Reviews. In *Proceedings of the 40th Annual Meeting on Association for Computational Linguistics*, 417–424. Stroudsburg, PA, USA: Association for Computational Linguistics.

Tweedie, F. J. and Baayen, R. H. 1998. How Variable May a Constant Be? Measures of Lexical Richness in Perspective. *Computers and the Humanities* 32 (5): 323–352.

Viswanathan, V., Mooney, R. and Ghosh, J. 2014. Detecting Useful Business Reviews using Stylometric Features. *cs.utexas.edu* .

Viviani, M. and Pasi, G. 2017. Quantifier Guided Aggregation for the Veracity Assessment of Online Reviews. *International Journal of Intelligent Systems* 32 (5): 481–501.

Vrij, A. 2008. *Detecting lies and deceit: pitfalls and opportunities*. New York, NY, US: Wiley-Interscience.

Wahyuni, E. D. and Djunaidy, A. 2016. Fake Review Detection From a Product Review Using Modified Method of Iterative Computation Framework. In *MATEC Web of Conferences*, 03003. The 3rd Bali International Seminar on Science and Technology (BISSTECH 2015).

Wang, D., Irani, D. and Pu, C. 2011a. A Social-spam Detection Framework. In *Proceedings of the 8th Annual Collaboration, Electronic messaging, Anti-Abuse and Spam Conference on - CEAS '11*, 46–54. New York, New York, USA: ACM Press.

Wang, G., Xie, S., Liu, B. and Yu, P. S. 2011b. Review Graph Based Online Store Review Spammer Detection. In *Proceedings of the 2011 IEEE 11th International Conference on Data Mining*, 1242–1247. Washington, DC, USA: IEEE Computer Society.

Wang, G., Xie, S., Liu, B. and Yu, P. S. 2012. Identify Online Store Review Spammers via Social Review Graph. *ACM Trans. Intell. Syst. Technol.* 3 (4): 61.

Wang, Q., Liang, B., Shi, W., Liang, Z. and Sun, W. 2010a. Detecting Spam Comments with Malicious Users' Behavioral Characteristics. In *2010 IEEE International Conference on Information Theory and Information Security*, 563–567.

Wang, W., Zeng, G. and Tang, D. 2010b. Using Evidence Based Content Trust Model for Spam Detection. *Expert Systems with Applications* 37 (8): 5599–5606.

Wang, X., Liu, K., He, S. and Zhao, J. 2016. Learning to Represent Review with Tensor Decomposition for Spam Detection. In *2016 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 866875.

Wang, X., Liu, K. and Zhao, J. 2017. Handling Cold-Start Problem in Review Spam Detection by Jointly Embedding Texts and Behaviors. In *proceedings of the Annual Meeting of the Association for Computational Linguistics (ACL-2017)*.

Wang, Z., Gu, S. and Xu, X. 2018. GSLDA: LDA-based group spamming detection in product reviews. *Applied Intelligence* .

Wiebe, J. 2000. Learning Subjective Adjectives from Corpora. In *Proceedings of the Seventeenth National Conference on Artificial Intelligence and Twelfth Conference on Innovative Applications of Artificial Intelligence*, 735–740. AAAI Press.

Witten, I. H. and Frank, E. 2005. *Data Mining: Practical Machine Learning Tools and Techniques, Second Edition (Morgan Kaufmann Series in Data Management Systems)*. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc.

Wu, G., Greene, D. and Cunningham, P. 2010a. Merging Multiple Criteria to Identify Suspicious Reviews. In *Proceedings of the fourth ACM conference on Recommender systems - RecSys '10*, 241. New York, New York, USA: ACM Press.

Wu, G., Greene, D., Smyth, B. and Cunningham, P. 2010b. Distortion as a Validation Criterion in the Identification of Suspicious Reviews. In *Proceedings of the First Workshop on Social Media Analytics - SOMA '10*, 10–13. New York, New York, USA: ACM Press.

Xie, S., Wang, G., Lin, S. and Yu, P. S. 2012a. Review Spam Detection via Temporal Pattern Discovery. In *Proceedings of the 21st international conference companion on World Wide Web*, 823–831. New York, NY, USA: ACM.

Xie, S., Wang, G., Lin, S. and Yu, P. S. 2012b. Review Spam Detection via Time Series Pattern Discovery. In *Proceedings of the 21st international conference companion on World Wide Web - WWW '12 Companion*, 635. New York, New York, USA: ACM Press.

Xu, C. and Zhang, J. 2015a. *Combating Product Review Spam Campaigns via Multiple Heterogeneous Pairwise Features*, Ch. 20, 172–180.

Xu, C. and Zhang, J. 2015b. Towards Collusive Fraud Detection in Online Reviews. In *Proceedings of the 2015 IEEE International Conference on Data Mining (ICDM)*, 1051–1056. Washington, DC, USA: IEEE Computer Society.

Xu, C., Zhang, J., Chang, K. and Long, C. 2013. Uncovering Collusive Spammers in Chinese Review Websites. In *Proceedings of the 22Nd ACM International Conference on Information & Knowledge Management*, 979–988. New York, NY, USA: ACM.

Xu, Y., Shi, B., Tian, W. and Lam, W. 2015. A Unified Model for Unsupervised Opinion Spamming Detection Incorporating Text Generality. In *Proceedings of the 24th International Conference on Artificial Intelligence*, 725–731. AAAI Press.

Xue, H., Li, F., Seo, H. and Pluretti, R. 2015. Trust-Aware Review Spam Detection. In *Trustcom/BigDataSE/ISPA, 2015 IEEE*, 726–733.

Yang, X. 2015. One methodology for spam review detection based on review coherence metrics. In *Intelligent Computing and Internet of Things (ICIT), 2014 International Conference on*, 99–102.

Yoo, K.-H. and Gretzel, U. 2009. Comparison of Deceptive and Truthful Travel Reviews. In *Information and Communication Technologies in Tourism 2009*, 37–47. The Netherlands: Springer Vienna.

Yule, G. U. 1939. On Sentence-Length as a Statistical Characteristic of Style in Prose: With Application to Two Cases of Disputed Authorship. *Biometrika* 30 (3/4): 363–390.

Yule, G. U. 1944. *The Statistical Study of Literary Vocabulary.*. Cambridge University Press.

Zakaluk, B. L. and Samuels, S. J. 1998. *Readability: Its Past, Present, and Future*. International Reading Association.

Zhang, W., Bu, C., Yoshida, T. and Zhang, S. 2016a. CoSpa: A Co-training Approach for Spam Review Identification with Support Vector Machine. *Information* 7 (1): 12.

Zhang, W., Jiang, Y. and Yoshida, T. 2016b. *Deep Context Identification of Deceptive Reviews Using Word Vectors*, 213–224. Singapore: Springer Singapore.

Zhang, W., Lau, R. Y. K. and Li, C. 2014. Adaptive Big Data Analytics for Deceptive Review Detection in Online Social Media. In *ICIS* (eds. M. D. Myers and D. W. Straub). Association for Information Systems.

Zheng, R., Li, J., Chen, H. and Huang, Z. 2006. A Framework for Authorship Identification of Online Messages: Writing-style Features and Classification Techniques. *Journal of the American Society for Information Science and Technology* 57 (3): 378–393.

Zhiyuli, A., Liang, X. and Wang, Y. 2015. Discerning the Trend: Concealing Deceptive Reviews. In *Systems, Man, and Cybernetics (SMC), 2015 IEEE International Conference on*, 1833–1838.

Zhou, D., Bousquet, O., Lal, T. N., Weston, J. and Schölkopf, B. 2003a. Learning with Local and Global Consistency. In *Proceedings of the 16th International Conference on Neural Information Processing Systems*, 321–328. Cambridge, MA, USA: MIT Press.

Zhou, D., Weston, J., Gretton, A., Bousquet, O. and Schölkopf, B. 2003b. Ranking on Data Manifolds. In *Proceedings of the 16th International Conference on Neural Information Processing Systems*, 169–176. Cambridge, MA, USA: MIT Press.