

Opinion Spam Detection: Detecting Fake Reviews and Reviewers

Many names: Spam Review, Fake Review, Bogus Review, Deceptive review
Opinion Spammer, Review Spammer, Fake Reviewer, Shill (Stooge or Plant),
(See this [The New York Times](#) front page article, Jan. 26, 2012)
(Bloomberg [BusinessWeek](#), Sept. 29, 2011 and [more ...](#))

New Book:  [Sentiment Analysis: mining opinions, sentiments, and emotions.](#)
Cambridge University Press, available from March 2015.

Book: [Sentiment Analysis and Opinion Mining](#) (Introduction and Survey), Morgan & Claypool, May 2012.

Introduction

It has become a common practice for people to read online opinions/reviews for different purposes. For example, if one wants to buy a product, one typically goes to a review site (e.g., amazon.com) to read some reviews of the product. If most reviews are positive, one is likely to buy the product. If most reviews are negative, one will almost certainly not buy it. Positive opinions can result in significant financial gains and/or fames for businesses, organizations and individuals. This, unfortunately, gives strong incentives for opinion spamming.

Can you figure out which of these [three reviews](#) are fake?

Opinion Spamming: It refers to "illegal" activities (e.g., *writing fake reviews*, also called *shilling*) that try to mislead readers or automated [opinion mining and sentiment analysis](#) systems by giving undeserving positive opinions to some target entities in order to promote the entities and/or by giving false negative opinions to some other entities in order to damage their reputations. Opinion spam has many forms, e.g., *fake reviews* (also called *bogus reviews*), *fake comments*, *fake blogs*, *fake social network postings*, *deceptions*, and *deceptive messages*.

We believe that as opinions on the Web are increasingly used in practice by consumers, organizations, and businesses for their decision making, opinion spamming will get worse and also more sophisticated. Detecting spam reviews or opinions will become more and more critical. The situation is already quite bad.

To the best of our knowledge, my group is the first to conduct research on detecting fake reviews and reviewers (or shills). Our first paper was published in 2007, and subsequent papers were published in 2008, 2010, and 2012. Both my books [Web Data Mining](#) and [Sentiment Analysis and Opinion Mining](#) discuss the issue.

NOTE: This is closely related to **Astroturfing**: "Astroturfing refers to political, advertising, or public relations campaigns that are designed to mask the sponsors of the message to give the appearance of coming from a disinterested, grassroots participant. Astroturfing is intended to give the statements the

credibility of an independent entity by withholding information about the source's financial connection. The term is a derivation of AstroTurf, a brand of synthetic carpeting designed to look like natural grass." Quoted from the [Wikipedia](#) page.

Acknowledgement: This project has been partially funded by *National Science Foundation, Microsoft, and Google*

Fake Review Detection

We have used supervised learning, pattern discovery, graph-based methods, and relational modeling to solve the problem. Below are some main signals that we have used:

1. Review content:
 1. **Lexical features** such as word n-grams, part-of-speech n-grams, and other lexical attributes.
 2. **Content and style similarity** of reviews from different reviewers.
 3. **Semantic inconsistency** (we have never used this kind of features). For example, a reviewer wrote "My wife and I bought this car ..." in one review and then in another review he/she wrote "My husband really love ..." (I heard this example from a friend in a company which actively detects fake reviews).
2. Reviewer abnormal behaviors:
 1. **Public data available from Web sites**, e.g., reviewer id, time of posting, frequency of posting, first reviews of products, and many more. For example, do you see anything wrong with the reviews from this user, [Big John](#)? What about after you see the reviews of these two users, [Cletus](#) and [Jake](#)? In fact, if you browse the reviews of their reviewed products, you will find another suspicious user/reviewer. This is just one example of atypical behaviors that our algorithms are able to discover.
 2. **Web site private/internal data** (we have not used such data, but they are extremely useful), e.g., IP and MAC addresses, time taking to post a review, physical location of the reviewer, etc (a lot of them).
3. Product related features: E.g., product description, sales volume, and sales rank
4. Relationships: Complex relationships among reviewers, reviews, and entities (e.g., products and stores).

Some Fake Review Cases in the News

- **Must Read:** [Samsung probed in Taiwan over 'fake web reviews'](#), April 16, 2013 [BBC News].
- **Must Read:** [Woman Paid To Post Five-Star Google Feedback](#), September 15, 2012 [ABC7 News] (our algorithm in [WWW-2012](#) paper can catch them).
- **Must Read:** [Are You Buying Reviews For Google Places?](#), Local Internet Marketing, January 27, 2012. (our algorithm in [WWW-2012](#) paper can catch them).
- **Must Read:** [The Best Book Reviews Money Can Buy](#), August 25, 2012 [The New York Times].
- **Must Read:** [For \\$2 a Star, an Online Retailer Gets 5-Star Product Reviews](#), Jan 26, 2012 [The New York Times].
- **Must Read:** [Amazon Glitch Unmasks War Of Reviewers](#), February 14, 2004 [The New York Times].
- [Charges Settled Over Fake Reviews on iTunes](#), August 26, 2010 [The New York Times].
- [Company Settles Case of Reviews It Faked](#), July 14, 2009 [The New York Times].
- [TripAdvisor warns consumers about fake reviews](#), July 16, 2009 [USA Today].
- [Belkin's Development Rep is Hiring People to Write Fake Positive Amazon Reviews](#), January 16,

2009 [The Daily Background].

- [A Fake Amazon Reviewer Confesses](#), July 9, 2009 [The Wall Street Journal].

Professional Fake Review Writing Services (some *Reputation Management* companies)

- [Post positive reviews](#) (this site's content and even URL keep changing)
- [Sponsored reviews](#) (a site where advertisers and bloggers get in touch to write paid reviews)
- [Pay per post](#) (also a site where advertisers and bloggers get in touch to write paid reviews)
- [Need someone to write positive reviews about our company](#) (budget: \$250-\$750 USD)
- [Fake review writer](#)
- [Product review writer for hire](#)
- [Hire a content writer](#)
- [Fake Amazon book reviews \(hiring book reviewers\)](#)
- [People are just having fun \(not serious\)](#)

How to Spot Fake Reviews Manually

- [30 Ways You Can Spot Fake Online Reviews -- Test Your Ability!](#)
- [3 Tips for Spotting Fake Product Reviews - From Someone Who Wrote Them](#)
- [Hone Your Eye for Fake Online Reviews](#)
- [How to spot fake user reviews](#)

I am doubtful that people can really spot fake reviews reliably (especially those well written ones). I have done experiments with 30+ students to show otherwise. One of the fallacies is that people usually think others would write like them or should write in certain ways.

Manipulating Social Media (sock puppets - fake identities - fake personas)

- [Revealed: US spy operation that manipulates social media](#), Guardian.co.uk, Thursday 17 March 2011.
- [America's absurd stab at systematising sock puppetry](#), Guardian.co.uk, Thursday 17 March 2011.

China's Internet "Water Army" (Shuijun) - Opinion Spammers

- You can hire people to write and post fake reviews or comments, and even bribe staff at review, forum and microblog sites to delete posts that you do not like.
- ['Water Army' Whistleblower Threatened](#), January 7, 2011, People's Daily.
- [The Chinese Online "Water Army"](#), June 25, 2010, Wired.com.
- If you read Chinese, see [this description](#) from Baidu Baike at baidu.com.

Data Sets

- [Amazon Product Review Data \(Huge\)](#) used in (Jindal and Liu, WWW-2007; WSDM-2008; Lim et al, CIKM-2010; Jindal, Liu and Lim, CIKM-2010; Mukherjee et al. WWW-2011; Mukherjee, Liu and Glance, WWW-2012) for review spam (fake review) detection. It has information about reviewers, review text, ratings, product info, etc. Due to the large file size, you may need to use *Download Accelerator Plus* (DAP) to download. If you use this data, please cite (Jindal and Liu,

WSDM-2008).

Publications

1. Jing Wang, Clement. T. Yu, Philip S. Yu, Bing Liu, Weiyi Meng. "Diversionary comments under blog posts." Accepted. *ACM Transactions on the Web (TWEB)*, 2015.
2. Huayi Li, Zhiyuan Chen, Arjun Mukherjee, Bing Liu and Jidong Shao. "Analyzing and Detecting Opinion Spam on a Large-scale Dataset via Temporal and Spatial Patterns." Short paper at *ICWSM-2015*, 2015.
3. Huayi Li, Arjun Mukherjee, Bing Liu, Rachel Kornfieldz and Sherry Emery. [Detecting Campaign Promoters on Twitter using Markov Random Fields](#). to appear in *Proceedings of IEEE International Conference on Data Mining (ICDM-2014)*, December 14-17, 2014.
4. Huayi Li, Zhiyuan Chen, Bing Liu, Xiaokai Wei and Jidong Shao. [Spotting Fake Reviews via Collective Positive-Unlabeled Learning](#). to appear in *Proceedings of IEEE International Conference on Data Mining (ICDM-2014, short paper)*, December 14-17, 2014.
5. Tieyun Qian, Bing Liu. [Identifying Multiple Userids of the Same Author](#). To appear in *Proceedings of Conference on Empirical Methods in Natural Language Processing (EMNLP-2013)*, October 18-21, 2013, Seattle, USA.
6. Arjun Mukherjee, Abhinav Kumar, Bing Liu, Junhui Wang, Meichun Hsu, Malu Castellanos, and Riddhiman Ghosh. [Spotting Opinion Spammers using Behavioral Footprints](#). To appear in *Proceedings of SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD-2013)*, August 11-14 2013 in Chicago, USA.
7. Arjun Mukherjee, Vivek Venkataraman, Bing Liu, and Natalie Glance. [What Yelp Fake Review Filter Might Be Doing](#). *Proceedings of The International AAAI Conference on Weblogs and Social Media (ICWSM-2013)*, July 8-10, 2013, Boston, USA.
8. Geli Fei, Arjun Mukherjee, Bing Liu, Meichun Hsu, Malu Castellanos, and Riddhiman Ghosh. [Exploiting Burstiness in Reviews for Review Spammer Detection](#). *Proceedings of The International AAAI Conference on Weblogs and Social Media (ICWSM-2013)*, July 8-10, 2013, Boston, USA.
9. Arjun Mukherjee, Bing Liu, and Natalie Glance. [Spotting Fake Reviewer Groups in Consumer Reviews](#). *International World Wide Web Conference (WWW-2012)*, Lyon, France, April 16-20, 2012. (See [media coverage](#) of this work from April 16, 2012)
10. Guan Wang, Sihong Xie, Bing Liu, Philip S. Yu. [Identify Online Store Review Spammers via Social Review Graph](#). *ACM Transactions on Intelligent Systems and Technology*, accepted for publication, 2011.
11. Guan Wang, Sihong Xie, Bing Liu, Philip S. Yu. [Review Graph based Online Store Review Spammer Detection](#). *ICDM-2011*, 2011.
12. Arjun Mukherjee, Bing Liu, Junhui Wang, Natalie Glance, Nitin Jindal. [Detecting Group Review Spam](#). *WWW-2011 poster paper*, 2011.

13. Nitin Jindal, Bing Liu and Ee-Peng Lim. ["Finding Unusual Review Patterns Using Unexpected Rules"](#) *Proceedings of the 19th ACM International Conference on Information and Knowledge Management (CIKM-2010, short paper)*, Toronto, Canada, Oct 26 - 30, 2010.
14. Ee-Peng Lim, Viet-An Nguyen, Nitin Jindal, Bing Liu and Hady Lauw. ["Detecting Product Review Spammers using Rating Behaviors."](#) *Proceedings of the 19th ACM International Conference on Information and Knowledge Management (CIKM-2010, full paper)*, Toronto, Canada, Oct 26 - 30, 2010.
15. Nitin Jindal and Bing Liu. ["Opinion Spam and Analysis."](#) *Proceedings of First ACM International Conference on Web Search and Data Mining (WSDM-2008)*, Feb 11-12, 2008, Stanford University, Stanford, California, USA.
16. Nitin Jindal and Bing Liu. ["Review Spam Detection."](#) *Proceedings of WWW-2007* (poster paper), May 8-12, Banff, Canada.

Three Reviews - Can you figure out which ones are fake?

1. I want to make this review in order to comment on the excellent service that my mother and I received on the Serenade of the Seas, a cruise line for Royal Caribbean. There was a lot of things to do in the morning and afternoon portion for the 7 days that we were on the ship. We went to 6 different islands and saw some amazing sites! It was definitely worth the effort of planning beforehand. The dinner service was 5 star for sure. One of our main waiters, Muhammad was one of the nicest people I have ever met. However, I am not one for clubbing, drinking, or gambling, so the nights were pretty slow for me because there was not much else to do. Either than that, I recommend the Serenade to anyone who is looking for excellent service, excellent food, and a week full of amazing day-activities!
2. This movie starring big names - Tom Hanks, Sandra Bullock, Viola Davis, and John Goodman - is one of the most emotionally endearing films of 2012. While some might argue that this film was "too Hollywood" and others might see the film solely because of the cast, it is Thomas Horn's performance as young Oskar that is deserving of awards. The story is about a 9-year-old boy on a journey to make sense of his father's tragic death in the 9/11 attacks on the World Trade Center. Oskar is a bright and nervous adventurer calmed only by the rattle of a tambourine in his ear. "I got tested once to see if I had Asperger's disease," the boy offers in explain of his odd behavior. "The tests weren't definitive." One year after the tragedy, Oskar finds a key in his father's closet and thus begins a quest to find the missing lock. Oskar's battle to control his emotional anxiety and form and mend relationships proves difficult, even with his mother. "If the sun were to explode, you wouldn't even know about it for eight minutes," Oskar narrates. "For eight minutes, the world would still be bright and it would still feel warm." Those fleeting eight minutes Oskar has left of his father make for two hours and nine minutes of Extremely Emotional and Incredibly Inspiring film. Leaving the theatre, emotionally drained, it is a wonder where a movie like this has been. We saw Fahrenheit 9/11 and United 93, but finally here is the story of a New York family's struggle to understand why on "the worst day" innocent people would die. I highly recommend this movie as a must see.
3. High Points: Guacamole burger was quite tall; clam chowder was tasty. The decor was pretty good, but not worth the downsides. Low Points: Noisy, noisy, noisy. The appetizers weren't very good at all. And the service kind of lagged. A cross between Las Vegas and Disney world, but on the cheesy side. This Cafe is a place where you eat inside a plastic rain forest. The walls are lined with

fake trees, plants, and wildlife, including animatronic animals. A flowing waterfall makes sure that you won't hear the conversations of your neighbors without yelling. I could see it being fun for a child's birthday party (there were several that occurred during our meal), but not a place to go if you're looking for a good meal.

4. **The answer is at the bottom of my homepage. If you get them right, please let me know your clues.** You can click my name below to get to my homepage.

Created by [Bing Liu](#), 2008.