Temporal and Sentimental Analysis of A Real Case of Fake Reviews in Taiwan

Chih-Chien Wang
National Taipei University
Taipei, Taiwan
wangson@mail.ntpu.edu.tw

Min-Yuh Day
Tamkang University
Taipei, Taiwan
myday@mail.tku.edu.tw

Chien-Chang Chen
Tamkang University
Taipei, Taiwan
ccchen34@mail.tku.edu.tw

Jai-Wei Liou
National Taipei University
Taipei, Taiwan
bgga1013344@gmail.com

***Abstract*—Product reviews are important information sources for consumers as they make their purchasing decisions. However, some unethical firms hire fake reviewers to generate biased positive reviews to promote their product and to damage the product reputations of their competitors. From the point of view of online product review platform providers, it is essential to keep the platform neutral and unbiased by detecting fake reviews and preventing fake reviewers from spreading biased reviews. In the current study, we attempt to use temporal and sentiment analyses as cues to separate fake reviews from authentic product reviews. Real case data of fake reviews in Taiwan was used for this temporal and sentiment analysis. Based on the analysis results, we find that fake reviewers usually generated and replied to fake reviewers during normal work hours. In contrast, ordinary users only generated and replied to a small proportion of normal product reviews during work hours. They generated and replied to normal product reviews the most during off-work hours and weekends. Additionally, the current study also revealed that more than half of fake reviewers replied others' responses to their own fake reviews no later than within one day. The research results revealed that temporal and sentiment analyses have the potential to serve as cues to detect fake reviews and fake reviewers.**

Keyword—Spammers; Fake Reviewers; Fake Review; Temporal Analysis; Sentiment Analysis

ASONAM '17, July 31-August 03, 2017, Sydney, Australia © 2017 Copyright is held by the owner/author(s). Publication rights licensed to ACM.
ACM ISBN 978-1-4503-4993-2/17/07…$15.00
https://doi.org/10.1145/3110025.3116206

I. INTRODUCTION

The rapid development of telecommunication technology and the high penetration rate of Internet access makes product review websites major information sources when consumers make their purchase decisions. People usually surf product review websites to seek others' opinions and consumption experiences before buying a product. After consuming the product, people may share their consumption experience by reviewing and scoring the products on product review websites.

Positive opinions lead people to buy the products, while negative opinions prevent consumers from buying products. Since online opinions are important information sources for consumers making purchase decisions, firms sometimes want to manipulate the opinions. Some unethical firms may hire full-time or part-time workers to create biased fake reviews to promote their products and demote their competitors' products. Some of the fake reviewers even create multiple accounts to post a significant amount of fake opinions to increase their influence.

While many consumers consult opinions shared on the online review websites before making a purchase decision, they may be deceived by fake reviews purposely provided by unethical companies. To avoid the damaging effect fake reviews have on the value of product review websites, academics and industry practices have devoted themselves to developing mechanisms to detect fake reviews and fake reviewers [1-8].

The existence of fake reviews is a serious issue for product review websites, since these fake reviews may damage the trustworthiness of the sites. People may be skeptical of online opinions when some of them are discovered to be fake rather than real opinions based on consumption experience. Product review websites may devote time and effort into preventing the spread of fake reviews. One famous case was the lawsuit by Amazon against more than 1,114 fake reviewers [9]. These fake reviewers mislead consumers using fake reviews, which affected Amazon.com's reputation. Consumers read the fake product reviews and some products even became bestsellers because of the fake reviews.

Previous studies have proposed some ideas for mechanisms that will detect fake reviews. For example, the posted reviews and their replies can be treated as a social network relationship, which can be cues for fake reviewer detection [5, 10, 11]. Sentiment analysis can be another approach to detect fake reviews [4, 12]. Also, the similarity of reviews can also be a cue for fake reviews.

Fake reviews are usually posted by part-time or full-time workers hired by companies. These workers post fake reviews in exchange for payment. Nevertheless, normal users post their opinions based on their consumption experience, whether positive or negative. The difference in the behavior of normal reviewers versus fake review reviewers has the potential to be used to detect the fake reviewers. Previous research had proposed the idea of using behavior analysis to detect fake reviewers [13-15].

The current study analyzes the behavior of both fake reviewers and normal users to discuss the possibility of using behavior analysis to detect fake reviewers. Temporal analysis and sentimental analysis of reviews were used to reveal the differences between fake and normal reviews. A real case of fake reviews in Taiwan was used in the current study.

II. Literature Reviews

There are four steps in studying fake reviews, as Fig. 1 reveals. First, researchers need to obtain a dataset or corpus containing both fake reviews and ordinary reviews. The collected fake reviews should be the "ground truth" so that researchers can use it to develop an anti-fake review mechanism. Then, the researchers propose an idea to find cues that separate the fake and ordinary product reviews. Also, they have to choose quantitative techniques or algorithms that use these cues to separate the fake reviews from the ordinary reviews. Finally, the researchers may

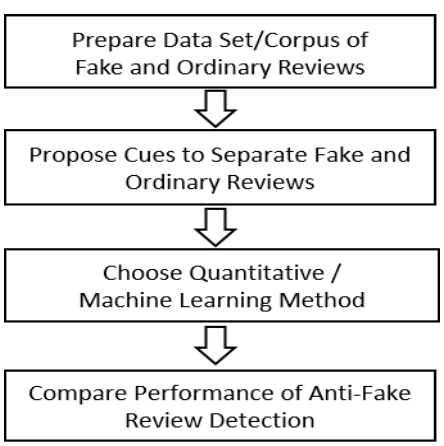


Figure 1. Research Procedure for Fake Review Detection

compare their performance for fake review detection.

The current study used a real case of fake reviews in Taiwan and conducted temporal and sentimental analyses to detect fake reviews and to verify the feasibility of using those analyses to detect fakes reviews.

A. *Fake Review Collection*

It is important to use a sound fake review dataset/corpus in the development of anti-fake mechanisms. However, it is not easy to get the "ground truth" of fake reviews. There are at least three approaches to obtain fake reviews: Mark reviews manually as fake reviews, hire writers to write fake reviews, and use the real fake review cases. Real fake review cases are useful but rarely available.

1) *Mark Reviews as Fake Reviews*

Jindal and Liu are pioneers in fake review detection [16, 17]. They crawled data from Amazon.com and argued that there were three types of review spam: untruthful opinions, reviews on brands only, and non-reviews. They confirmed the existence of the proposed three kinds of review spam by detecting duplicate reviews and manually marking reviews as spam. The dataset was used to develop an anti-fake review detection approach [13, 18].

Jindal and Liu detected duplicate reviews and treated them as fake reviews. However, not all fake reviews are duplicate reviews. Fake reviewers may write "new" fake reviews rather than duplicate previous fake reviews as new fake reviews. If the fake review does not belong to the duplicate review, then it would not be included in the fake review list as designed by Jindal and Liu.

Researchers may also hire experts or normal users to mark reviews manually as fake reviews. This is another approach to get a fake review list. However, this approach is based on the assumption that these experts or normal users know can judge if a review is fake or ordinary consumption experience sharing.

2) *Hire Writers to write fake reviews*

Ott et al. proposed another way to get fake reviews: hiring writers from Amazon's Mechanical Turk to write fake reviews [19]. The hired workers were asked to write fake (deceptive) reviews for hotels. These fake reviews were mixed with ordinary reviews from TripAdvisor.com to formulate a database for analysis.

B. *Cues to Detect Fake Review*

Previous research has proposed different cues and ideas for detecting fake reviews. Review similarity, temporal analysis, and sentiment analysis are three approaches for detecting fake reviews.

1) *Similarity Analysis*

Similarity analysis has been a common approach for detecting fake reviews in previous fake review research. This approach is based on the assumption that fake reviews are repetitions of another fake reviews.

Algur et al. [20] used cosine similarity to detect fake reviews, by considering repeated and close to repetitive

TABLE 1. COLLECTED FAKE AND NORMAL REVIEWS

	Fake Reviews	Normal Reviews	Total
Post	*458*	*8363*	*8821*
Reply	*5245*	*111065*	*116310*

The field of *"post-id", "author", "date-and-time", "post-or-reply", "content"* are used in the current research.

reviews as spam reviews, and unique reviews as non-spam (ordinary) reviews. The experiment found that similarity could detect a large number of spam comments, eliminate spam reviews and retain unique reviews so that consumers can make the right purchase decisions.

Jindal and Liu [17] artificially marked the repetition of word of mouth and used Jaccard similarity to judge if a review was real or fake.

2) Temporal Analysis

Temporal analysis refers to a series of events or data changes over time, according to the order of occurrence. Since data or events change regularly and gradually with time, temporal analysis can be used to predict the direction of the phenomenon and its quantity.

The purpose of temporal analysis is to observe and analyze past data or events to predict future developments. Xie et al. [21] argued that temporal anomalies of fake reviews can be used to detect fake reviews.

For fake review spammers, it is part of their work to post fake reviews and respond to that review. Therefore, fake review spammers may post and reply reviews during their work hours. For these spammers, posting fake reviews is only a job. During their off-work hours, they will not post or reply to the reviews anymore. However, for normal users, sharing their consumption experience may be a part of their leisure activities. The behavior patterns are therefore different between fake review spammers and normal users.

3) Sentiment Analysis

Sentiment analysis can discover the attitude of the review author based on the text of the review. Sentiment polarity is a continuity value from positive to neutral to negative. Nevertheless, the continuity value of sentiment polarity can also be simplified into three categories: negative, neutral and positive.

Da Silva et al. [22] analyzed Twitter posts using sentiment analysis. They classified posts into positive and negative ones and argued that posts with neutral sentiment were more common compared to positive and negative ones. However, some Twitter posts were too short, which limited the function of sentiment analysis.

Researchers can hire raters to conduct sentiment analysis for online product reviews. However, the amount of online product reviews are so large so that manual markup will cost a significant amount of time and research budget. It is not practical to conduct a sentiment analysis by manual markup. Lau et al. [23] used sentiment lexicons to conduct an automated sentiment analysis of the review contents to detect spam reviews. WordNet, an English dictionary based on meaning of words, was used in their study to detect untruthful reviews. Their research revealed that text mining on semantic language modeling is a feasible approach for detecting spam reviews.

III. A Real Case of Fake Reviews in Taiwan

As mentioned above, identifying fake reviews is a challenge. Previous studies (such as Jindal and Liu [16] and [17]) have used similarity checks and/or manual markups to annotate "the ground truth" of fake reviews. In this kind of spam review dataset, duplicates and near-duplicates were considered as spam reviews. This approach was based on the assumption that fake reviews are similar and people can easily identity if a review is fake. However, review spammers may pretend that they are ordinary users. They change the content of their reviews to cheat users, specifically so users cannot precisely identify the fake reviews if the reviews are rewritten and look like normal consumption experience sharing.

A real case of opinion spam in Taiwan can be used in studying fake research [10, 11]. This data is a real case of fake reviews that occurred in Taiwan during 2012 to 2013. A large international mobile phone company based in Korea hired full time or part time employees to pretend they were customers writing product reviews in Taiwan in an attempt to influence the public opinion of their product and enhance their brand image. The fake reviews appeared on a large product review website in Taiwan that focuses on mobile phones.

In April 2013, a hacker posted several confidential marketing campaign documents of the company in the website (the website was not available after 2017). In the confidential documents, people received detailed information about how the company and its child company hired full-time and part-time workers to write fake reviews to promote their product. Such marketing malpractice violated Taiwan's Fair Trade Law. The company was fined by the Fair Trade Commission (FTC) in Taiwan and the company did not deny the truth of the leaked confidential documents.

This was the first time that research was able to get the "ground truth" of fake reviews, and people realized the existence of fake review. The dataset of this real case can now be used to explore the behavior of fake reviewers. Thus, the study adopted this dataset to conduct temporal and sentiment analysis.

IV. Methodology

A. Data collection

The data used in this study was the same as in [10, 11]. We focus on the both original posts and replies to the posts. As mentioned in TABLE 1, we collected 8,821 posts, including 458 fake review posts and 8,363 normal (non-spam) posts, respectively. Also, we got 116,310 replies, including 111,065 replies to the normal posts and 5,245 replies to the fake review posts. The numbers were slightly different from the origin dataset by [10, 11] since we found some reviews that were already moved to a junk folder of the product review website.

B. Procedure

The study used temporal and sentimental analysis to explore the post behavior of fake reviewers and normal reviewers.

1) Temporal Analysis

In the temporal analysis, we analyzed the post time by hours and by days of the week. In addition, we calculated the time duration between the original post and the first reply to the original post.

2) Automated Sentimental Analysis

The first step for sentiment analysis is Chinese word segmentation. We used python and Jieba segmentation tools, and open source segmentation tools (https://github.com/ldkrsi/jieba-zh_TW) for the purpose of segmenting sentences in traditional Chinese.

After segmenting a sentence in Chinese, we used a sentimental directory to count sentimental words. NTU Sentiment Dictionary (NTUSD), developed by National Taiwan University [24], was used in this step. NTUSD contains Chinese positive words and negative words.

The next step is to calculate the sentiment scores. We counted the positive and negative words that appeared in the fake and normal reviews. We used the following sentimental score formula to calculate the sentiment of reviews:

$$Sentiment\ \ Score = \frac{P - N}{T}$$

$$where$$

$$P = Count\ \ of\ \ Positvie\ \ Words$$

$$N = Count\ \ of\ \ Negative\ \ Words$$

$$\mathrm{T = Total\ \ Word\ \ Except\ \ Stop\ \ Words}$$

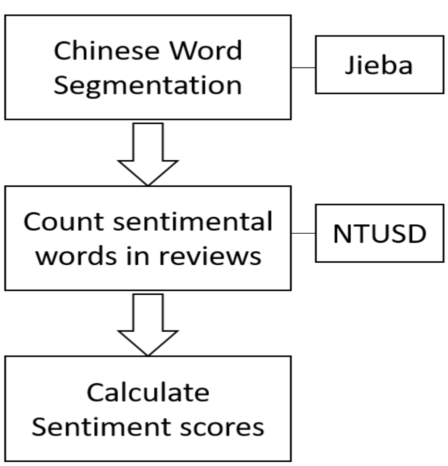


Figure 2. Automated Sentiment Analysis Procedure

The sentiment score calculates the proportion of positive and negative words in the whole product review. In this paper, the review was identified as positive if the sentimental score of the review was greater than zero, neutral if the score equaled zero, and as negative if the sentimental score was smaller than zero.

3) Manual Sentimental Analysis

To ensure that the results of the automated sentimental analysis were appropriate, we randomly selected 100 product reviews and asked raters to manually analyze these reviews. Raters were asked to rate the sentiment of the product review using an 11-point scale, ranging from -5 (negative) to 0 (neutral) to 5 (positive). We used a correlation coefficient to compared artificial (manually rate) sentiment scores and automated sentiment scores.

V. Results

A. Temporal Analysis of Posts and Replies

1) Temporal Analysis by Hour

Fig. 3 shows the trend of original posts of both spam posts and normal posts by hour of day. We can see that peak hours of fake review posts are between 9 a.m. and 6 p.m. This period is typically office hours between 9 a.m. and 6 p.m. Normal posts, however, gradually increased after 6 p.m. and reached their peak at 11 p.m. The period between 6 p.m.

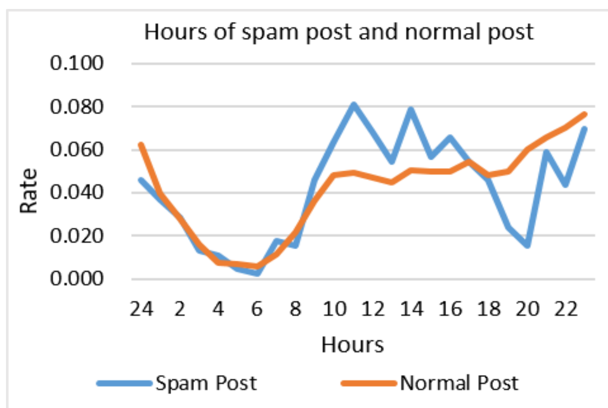


Figure 3. Hours of spam post and normal post

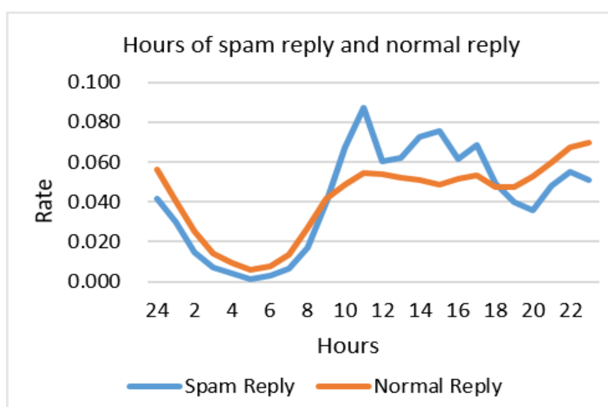


Figure 4. Hours of spam reply and normal reply

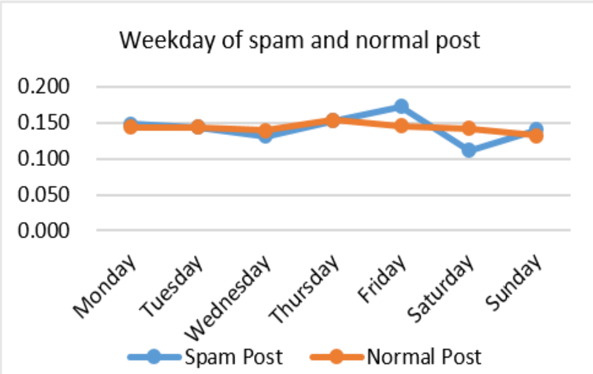


Figure 5. Weekday of spam post and normal post

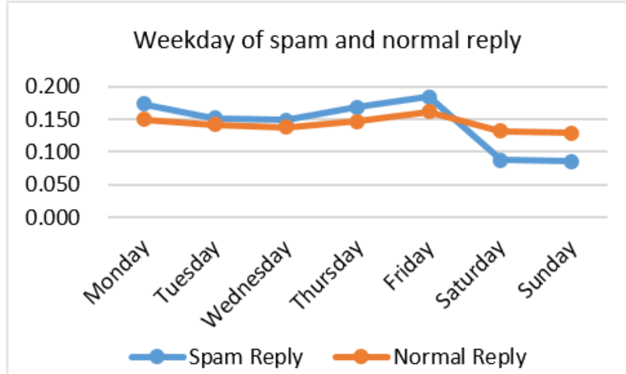


Figure. 6. Weekday of spam reply and normal reply

to 11 p.m. are typical off-work hours. After 12 o'clock midnight, the percentage of both fake review posts and normal posts significantly declined. Nevertheless, the low percentage of posted reviews between 1 a.m. and 8 a.m. makes sense, since people, whether fake reviewers or ordinary users, usually rest between 1 a.m. and 8 a.m.

Fig. 4 shows the trend of replies to original posts by both fake reviewers and normal reviewers by hour of day. The pattern of the time distribution of replies is similar to that of posts, as shown in Fig. 3.

Based on the results of Fig. 3 and Fig. 4, it is a reasonable inference that fake reviewers (spammers) usually post and reply during working hours only. Thus, post time can serve as a cue for detecting fake reviews.

2) Temporal Analysis by Day of the Week

Fig. 5 and Fig. 6 show a temporal analysis of post and reply rates by day of the week. As Fig. 5 and Fig. 6 reveal, the percentage of normal posts and replies are not obviously different between weekdays and weekends. In contrast, the percentage of spam posts and spam replies slightly decreases on weekends. The decrease in percentage of fake replies on the weekend is more obvious than that of fake posts on the weekend. The difference, however, is not large enough to serve as a major cue for detecting fake reviews.

Based on the results of Fig. 5 and Fig. 6, it is a reasonable inference that fake reviewers (spammers) usually post and reply on weekdays. Thus, post day can serve as a cue to detect fake reviews. Nevertheless, some fake reviews may still be posted on weekends. The post day cannot be used alone and should be used with other cues to detect fake reviews.

3) Duration between Original Post and First reply

When a product review is posted, others may reply to the original post. The discussion between original posters and other repliers may influence the original post. Thus, fake reviewers do not only post original posts, but they also reply to others' posts. We are curious if the fake reviewers will reply to posts of other fake reviews to help formulate the majority opinion for the original posts.

TABLE 2 shows the time distribution of the self-reply amount and percentage by hours. The self-reply as defined here is a replies to fake reviews by fake the reviewer themselves or other fake reviewers. The second column of TABLE 2 represents the number of replies by fake reviewers themselves or other fake reviewers. The third column represents the accumulated percentage that fake reviewers self-replied. The fourth column represents the accumulated amount of self-reply times. Of the 458 fake reviews collected in the current study, 22.3% self-replied within one hour, and half of them self-replied within 10 hours, by fake reviewers themselves or by other fake reviewers. Of the fake reviews, 62.4% were replied to by fake reviewers themselves or other fake reviewers.

Based on the results of TABLE 2, we can conclude that

TABLE 2. THE SPAM REPLY SITUATION

	Fake Reviews Replies by fake reviewers		
Reply in	**Self-reply amount**	**Self-reply percentage (%)**	**accumulate reply times**
1 hour	*102*	*22.3%*	*157*
2 hours	*136*	*29.7%*	*258*
3 hours	*168*	*36.7%*	*332*
4 hours	*185*	*40.4%*	*383*
5 hours	*195*	*42.6%*	*412*
6 hours	*201*	*43.9%*	*431*
7 hours	*209*	*45.6%*	*448*
8 hours	*212*	*46.3%*	*466*
9 hours	*216*	*47.2%*	*486*
10 hours	*224*	*48.9%*	*502*
11 hours	*231*	*50.4%*	*518*
12 hours	*235*	*51.3%*	*532*
13 hours	*239*	*52.2%*	*555*
14 hours	*240*	*52.4%*	*571*
15 hours	*241*	*52.6%*	*585*
16 hours	*243*	*53.1%*	*600*
17 hours	*246*	*53.7%*	*611*
18 hours	*250*	*54.6%*	*622*
19 hours	*252*	*55.0%*	*626*
20 hours	*252*	*55.0%*	*635*
21 hours	*254*	*55.5%*	*656*
22 hours	*257*	*56.1%*	*672*
23 hours	*260*	*56.8%*	*686*
24 hours	*260*	*56.8%*	*711*
after 24 hours	*286*	*62.4%*	*1010*

Note: the self-reply defined in this table is fake reviews replied to by the fake reviewers themselves or other fake reviewers.

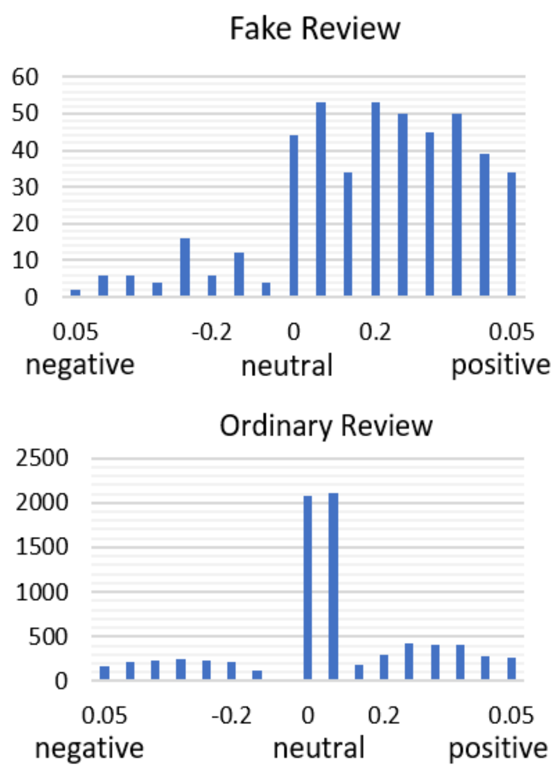
fake reviewers will reply to other fake reviewers. These results revealed that the relationship between posters and repliers could be a cue for detecting fake reviews.

B. Sentiment analysis

Fig. 7 reveals the sentiment comparison between fake reviews and ordinary reviews. We can find that most fake reviews are positive reviews. The proportion of positive fake reviews was larger than that of negative fake reviews. However, for ordinary reviews, most reviews were neutral. Only a tiny proportion of ordinary reviews were extremely positive or extremely negative. Based on Fig. 7, we found that most fake reviews had a positive sentiment, while most ordinary reviews had neutral or close to neutral sentiments.

TABLE 3 reveals the automated sentimental analysis results. As TABLE 3 reveals, most fake reviews posted (71.8%) were positive, while 18.5% of fake reviews were negative. However, only 38.6% of normal reviews were positive, while another 34.8% of normal reviews were negative. This result reveals that most fake reviews are positive reviews to promote products. Nevertheless, positive normal reviews by normal reviewers did not overwhelm with excess negative normal reviews. Thus, sentimental analysis can be a cue for detecting fake reviews.

TABLE 3 also reveals the sentimental analysis results of replies. The percentage of positive replies (40.2%) by fake reviewers is larger than negative ones (28.0%). However,



Note: The horizontal axis is sentimental score. The vertical axis is number of reviews.

Figure 7. Comparison sentiment of fake review and ordinary review

TABLE 3. SENTIMENT CALCULATED

	Fake Reviewers	Normal Reviewers
Original Post	*Positive: 329(71.8%)*	*Positive:3,092(38.6%)*
	Neutral:44(9.6%)	*Neutral: 2,082(26.5%)*
	Negative:85(18.5%)	*Negative: 2,731(34.8%)*
Reply to Post	*Positive: 2,109(40.2%)*	*Positive: 35,288(33.3%)*
	Neutral:1,666(31.7%)	*Neutral: 35,996(34.0%)*
	Negative:1,470(28.0%)	*Negative:34,536(32.6%)*

unlike original reviews, the percentage of positive replies by fake reviewers did not overwhelm the negative replies. Thus, the effect of using a sentimental analysis of replies to detect fake reviews was not as good as using sentimental analysis of original reviews.

TABLE 3 is based on an automated sentimental analysis using the assumption that automated sentimental analysis can discover the real sentiment of the product reviews. To confirm whether the automated sentimental analysis and human rates got similar results, the current study manually marked the sentiment of product reviews. Manual marking for sentimental analysis is time-consuming. Thus, we only randomly selected 100 product reviews to compare. The current study invited two raters to rate the sentiment of the product reviews manually.

To confirm whether the automated sentimental analysis results and human rating results were similar, we calculated the accuracy, Cohen's Kappa coefficient, and correlation coefficient between the two rating approaches. Based on TABLE 4, the Cohen's Kappa coefficient between human raters and the computer was only 0.13, which is unacceptable. If we omit the neutral sentiment reviews data, the Cohen's Kappa coefficient improves to 0.28, which is still low.

The Cohen's Kappa coefficient between the human rater 1 and human rate 2 is 0.46. If we remove the neutral sentimental reviews, the Cohen's Kappa coefficient increases to 0.81, which is acceptable. The correlation coefficient between human raters and the computer was only 0.23, which is low, modestly correlated and unacceptable. The correlation coefficient between human rater 1 and human rate 2 is 0.71, which is in the highly correlative level [25].

Based on the evidence of accuracy, Cohen's Kappa coefficient and the correlation coefficient, we should question the accuracy of the automated sentimental analysis

TABLE 4. COMPARISON OF MANUAL AND AUTOMATIC RATING

	human raters and computer	human rater 1 and human rate 2
Cohen's Kappa and Accuracy	*Kappa:0.13*	*Kappa:0.46*
	Accuracy:0.57	*Accuracy:0.68*
Cohen's Kappa and Accuracy without neutral data	*Kappa:0.28*	*Kappa:0.81*
	Accuracy:0.72	*Accuracy:0.92*
Correlation coefficient	*Correlation:0.23*	*Correlation:0.71*

used in the sentimental analysis for product reviews. At least, in the current research, the results of automatic sentimental analysis and human rating sentimental analysis was not highly correlative. Future studies should improve the automated sentimental analysis procedure to improve the analysis results.

VI. CONCLUSIONS

Fake reviewers provide misleading consumption experiences in product reviews to promote products and damage the reputation of competitors, which results in consumer misjudgment. If fake reviews are omnipresent in online product reviews, people may make the wrong decision because of these fake reviews.

From the temporal analysis results, we found that fake review spammers usually post fake reviews during work hours and workdays. Thus, post time and post day may serve as cues for anti-fake reviews. The current study also revealed that more than half of fake reviews received replies from other fake reviews. Thus, the social network between fake review posters and repliers can be cues to detect fake reviewers.

Besides, the current study also found that fake review spammers usually post product reviews with positive sentiments. However, normal reviewers post both positive and negative reviews. Thus, sentiments can be cues to detecting fake reviews.

The current study has several contributions. First, we used a real fake review case and adopted temporal analysis to reveal the behavior habits of the fake review spammers. Besides, we explored the possibility of using sentimental analysis to detect fake reviews.

This current study has several limitations. First, we found that it is not an easy task to analyze the sentiment of product reviews, since the sentiment of Chinese sentences is complicated. Also, the amount of product review datasets is large. Thus, it is not feasible to conduct manual sentimental analysis. The accuracy of automated sentimental analysis can hopefully increase so that future studies can detect fake reviews using automated sentimental analysis.

ACKNOWLEDGE

We would like to acknowledge MOST-Taiwan (Ministry of Science and Technology, Taiwan) for providing financial support to this research (MOST 105-2410-H-305 -065 - MY2).

REFERENCE

[1] L. Akoglu, H. Tong, and D. Koutra, "Graph based anomaly detection and description: a survey," Data Mining and Knowledge Discovery, vol. 29, pp. 626-688, May 2015 2015.

[2] Y.-R. Chen and H.-H. Chen, "Opinion Spammer Detection in Web Forum: A Real Case Study," in Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval, 2015, pp. 759-762.

[3] A. Colantonio, R. Di Pietro, M. Petrocchi, and A. Spognardi, "VISIO: A Visual Approach for Singularity Detection in Recommendation Systems," Trust, Privacy and Security in Digital Business, pp. 33-47, 2015.

[4] X. Deng and R. Chen, "Sentiment Analysis Based Online Restaurants Fake Reviews Hype Detection," Web Technologies and Applications, pp. 1-10, 2014.

[5] X. Hu, "Mining Content and Relations for Social Spammer Detection," 3700767 Ph.D., Arizona State University, Ann Arbor, 2015.

[6] J. Li, C. Cardie, and S. Li, "TopicSpam: a Topic-Model based approach for spam detection," ACL (2), pp. 217-221, 2013.

[7] A. Mukherjee and V. Venkataraman, "Opinion Spam Detection: An Unsupervised Approach using Generative Models," 2014.

[8] V. L. Rubin, Y. Chen, and N. Conroy, "Deception Detection for News: Three Types of Fakes," in The Proceedings of the Association for Information Science and Technology Annual Meeting (ASIST2015), 2015, pp. 6-10.

[9] B. Tuttle. (2015, May 15). Amazon Lawsuit Shows That Fake Online Reviews Are a Big Problem. Available: http://time.com/money/4078632/amazon-fake-online-reviews/

[10] C.-C. Wang, M.-Y. Day, and Y.-R. Lin, "Toward understanding the cliques of opinion spammers with social network analysis," in Advances in Social Networks Analysis and Mining (ASONAM), 2016 IEEE/ACM International Conference on, 2016, pp. 1163-1169.

[11] C.-C. Wang, M.-Y. Day, and Y.-R. Lin, "A Real Case Analytics on Social Network of Opinion Spammers," in Information Reuse and Integration (IRI), 2016 IEEE 17th International Conference on, 2016, pp. 623-630.

[12] R. Y. Chen, J. Y. Guo, and X. L. Deng, "Detecting fake reviews of hype about restaurants by sentiment analysis," Web-Age Information Management, pp. 22-30, 2014.

[13] N. Jindal, "Review spam and reviewer behavior analysis," 3431235 Ph.D., University of Illinois at Chicago, Ann Arbor, 2010.

[14] J. Chengzhang and D.-K. Kang, "Detecting spamming stores by analyzing their suspicious behaviors," in Advanced Communication Technology (ICACT), 2015 17th International Conference on, 2015, pp. 502-507.

[15] A. Mukherjee, A. Kumar, B. Liu, J. Wang, M. Hsu, M. Castellanos, et al., "Spotting opinion spammers using behavioral footprints," in Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining, 2013, pp. 632-640.

[16] N. Jindal and B. Liu, "Analyzing and Detecting Review Spam," in Data Mining, 2007. ICDM 2007. Seventh IEEE International Conference on, 2007, pp. 547-552.

[17] N. Jindal and B. Liu, "Opinion spam and analysis," in Proceedings of the 2008 International Conference on Web Search and Data Mining, 2008, pp. 219-230.

[18] A. Mukherjee, B. Liu, J. Wang, N. Glance, and N. Jindal, "Detecting group review spam," in Proceedings of the 20th international conference companion on World wide web, 2011, pp. 93-94.

[19] M. Ott, Y. Choi, C. Cardie, and J. T. Hancock, "Finding deceptive opinion spam by any stretch of the imagination," in Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies-Volume 1, 2011, pp. 309-319.

[20] S. P. Algur, A. P. Patil, P. Hiremath, and S. Shivashankar, "Conceptual level similarity measure based review spam detection," in Signal and Image Processing (ICSIP), 2010 International Conference on, 2010, pp. 416-423.

[21] S. Xie, G. Wang, S. Lin, and P. S. Yu, "Review spam detection via temporal pattern discovery," in Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining, 2012, pp. 823-831.

[22] N. F. Da Silva, E. R. Hruschka, and E. R. Hruschka, "Tweet sentiment analysis with classifier ensembles," Decision Support Systems, vol. 66, pp. 170-179, 2014.

[23] R. Y. Lau, S. Liao, R. C.-W. Kwok, K. Xu, Y. Xia, and Y. Li, "Text mining and probabilistic language modeling for online review spam detection," ACM Transactions on Management Information Systems (TMIS), vol. 2, p. 25, 2011.

[24] L. W. Ku and H. H. Chen, "Mining opinions from the Web: Beyond relevance retrieval," Journal of the American Society for Information Science and Technology, vol. 58, pp. 1838-1850, 2007.

[25] J. R. Landis and G. G. Koch, "The measurement of observer agreement for categorical data," biometrics, pp. 159-174, 1977.