Classification of Fake Product Ratings Using a Timeline Based Approach

Neha Thomas
GEMS Our Own English High School
Al Warqa'a, Dubai
nehasthomas@gmail.com

Dr. Susan Elias
School of Electronics Engineering
VIT University Chennai Campus, India
susan.elias@vit.ac.in

Abstract— **Detection of fake review and reviewers is currently a challenging problem in cyber space. It is challenging primarily due to the dynamic nature of the methodology used to fake the review. There are several aspects to be considered when analyzing reviews to classify them effective into genuine and fake. Sentiment analysis, opinion mining and intend mining are fields of research that try to accomplish the goal through Natural Language Processing of the text content of the review. In this paper, an approach that uses the review ratings evaluated along a timeline is presented. An Amazon dataset comprising of ratings indicated for a wide range of products was used for the analysis presented here. The analysis of the ratings was carried out for an electronic product over a period of six years. The computed average rating helps to identify linear classifiers that define solution boundaries within the dataspace. This enables a product specific classification of review ratings and suitable recommendations can also be generated automatically. The paper explains a methodology to evaluate the average product ratings over time and presents the research outcomes using a novel classification tool. The proposed approach helps to determine the optimal point to distinguish between fake and genuine ratings for each product.**

Index Terms: **Fake reviews, Fake Ratings, Product Ratings, Online Shopping, Amazon Dataset.**

I. INTRODUCTION

In an expatriate and tourist populated place like Dubai, online shopping has been growing fast. Almost everything you shop for on land can be accessed through your computer, and online shopping makes that extremely convenient. In Dubai, United Arab Emirates – also known as the "shopper's paradise" – the internet offers customers a plethora of online shopping options – which are, of course, made reliable with reviews by customers that have already bought the product. As genuine as these reviews may seem, they are fast becoming difficult to trust. A company may, for example, claim to have thousands of honest – but glowing – reviews about the products on their website, to make shoppers assured of their authenticity, but in reality, be paying people to post their polished reviews as though they were impartial. The dangers of these fake reviews include the common public being influenced by biased descriptions of products and creating an unfair and undependable environment between consumer and producer.

Another example could be competitors entering false and negative reviews about a product to dissuade the customer from buying it – thus keeping the shopper from a perfectly functioning product. The need for unbiased, reliable reviews and an identification process for detecting fake reviews has therefore become important. The problem is moderately realized. *Fakespot*, a tool to detect whether a product has reliable reviews or not, helps customers at Amazon.com – the world's largest online shopping portal – to be sure of the dependability of the product they buy.

The site analyzes the reviews on a particular product and then proceeds to give it a grade, telling you whether to use those reviews or not. This tool therefore allows you to see undistorted reviews with substantial accuracy. Comparing reviews with other genuine ones is a good aspect for detection of fake reviews. Language, vocabulary and sentence formation used in honest reviews sometimes contrast with spams/bogus reviews and hence create a good platform for assessment. Understanding the product, analyzing its uses, and probing for risks are other factors to consider when detecting fake reviews. Perceptions of fake and honest reviews provide a great podium to customers and consequently make their shopping lives easier and sham-free.

II. EXISITING TECHNIQUES AND MOTIVATION

An existing work [3, 4, 5] uses already created fake reviews generated by Amazon Mechanical Turk (AMT), based on a Yelp Dataset. The project analyses the fakeness of these genuine *fake* reviews and provides numerical information on how accurate these reviews are. It then computes the fakeness on *actual* reviews collected on popular websites such as TripAdvisor, etc. (also generated by AMT) and compares it with the results collected on the purposely-generated fake reviews. The project was a detailed study and used a novel method to reveal the difference between the accuracy of the two different sets of reviews and was therefore used to draw a conclusion. Several related research papers [5,6] also focus on addressing the challenges in detecting fake online reviews using other interesting techniques.

The motivation in carrying out the work presented in this paper was derived from the growing trend in faking reviews in both positive and negative aspects of the product. Online purchase is evolving to be a very convenient and cost effective approach to shop. But fake reviews and ratings make product identification very complicated and confusing. Hence a mechanism to identify fake reviews and reviewers are required to stop this online menace.

IJBAMR

III.PROPOSED STRATEGY FOR REVIEW RATING ANALYSIS

The proposed strategy to classify fake and genuine review ratings is presented in this section. The three primary factors used in our research include:

Factor1 (F1): Time span between the first and last review.
Factor2 (F2): Average review rating over the time span.
Factor3 (F3): Number of review ratings.

The methodology used to compute these factors from the dataset is presented as Algorithm 1 in the following section. Using factors F1 and F2, data points were created on a 2D space. Linear classifiers were determined to segregate the data points into solution areas for further analysis. Three classifiers were used to create six regions of interest. One classifier is parallel to the X Axis and the other two are parallel to the Y Axis. The values of the review ratings range from 1 to 5: one being the lowest and 5 the highest. The value of 3 was the *classification point* above which all reviews were positive and below which all reviews were negative. Hence our horizontal classifier is represented by the equation

$$Y = 3 \quad (1)$$

The position of the vertical classifier depends on several factors that vary according to the product. For instance – for electronic products, a popularity shelf life exists, which is defined by the dynamic nature of the market, competition, and other factors. Assuming the position 'P' is the beginning of the popularity period and 'S' is the ending of the popularity period, the vertical classifiers are defined by the equations that act as left and right boundaries:

$$X = P \quad (2)$$
$$X = S \quad (3)$$

The horizontal and two vertical classifiers thereby create six regions of interest, as given below:

Region 1: *Valid Positive Ratings*
Pre-popularity period + Positive reviews
Region 2: *Unacceptable Negative Ratings*
Pre-popularity period + Negative reviews
Region 3: *Acceptable Positive Ratings*
Popularity period + Positive reviews
Region 4: *Valid Negative Ratings*
Popularity period + Negative reviews
Region 5: *Consistent Positive Ratings*
Post-popularity period + Positive reviews
Region 6: *Consistent Negative Ratings*
Post-popularity period + Negative reviews

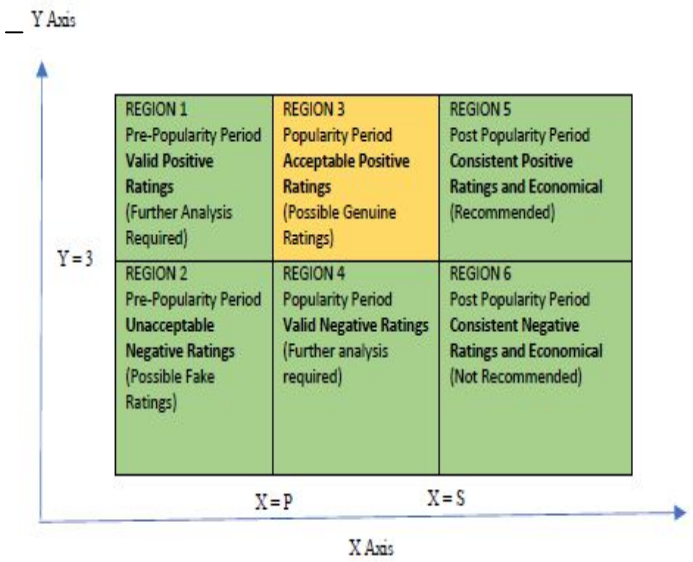


Fig. 1: Linear Classifiers and Solution Regions

If the product under consideration is genuinely good, we can say that

- Region 3 is the most acceptable choice for purchase because it has positive reviews throughout its popularity period.
- Region 2 is least acceptable and has *possibility of fake reviews* because it had rapid bad reviews at the start of its popularity period.
- Regions 1 and 4 contain valid ratings that can be used for decision making: Region 1 has positive reviews at the start of its popularity and Region 4 has bad reviews throughout its popularity period. The review content could be used for deeper analysis in this case.
- Regions 5 and 6 contain consistent ratings that can be used for decision making: Region 5 has positive reviews consistently and has a possibility of having associated discount offers and is therefore recommended. Region 6 has bad reviews consistently and is therefore not recommended.

The above linear classifiers are the fake rating detection filters proposed in this paper. The value assigned to the vertical classifiers are variable and differs for each product. The value that is suitable for 'P' and 'S' can be determined after analyzing the market trends for the product. The above six regions can vary in area depending in the position of the linear classifiers / fake rating detection filters.

The above methodology helps in quickly classifying the product into one of the 6 categories based on the evolved time. The number of review ratings (factor F3) could be used to further analyse the impact of the review ratings.

IV. DATA ANALYSIS

The Amazon Product Dataset was used for the research analysis presented in this subsection. The dataset was obtained from SNAP - Stanford Network Analysis Platform, a general-purpose network analysis and graph mining library [1]. A new

IJBAMR

and improved Amazon Dataset with redundant values removed was downloaded [2].

This dataset contains reviews of products – which includes text, ratings and helpfulness votes, links, and also provides information about the product metadata for 142.8 million reviews spanning May 1996 – July 2014. The reviews ranged from wide variety of products, such as books, movies, home and kitchen, video games, health and personal care, digital music, etc. The dataset used for this research was a "ratings only" dataset for Electronics Products, and contained no text reviews or metadata. The format of the dataset included the Reviewer ID, Product ID, Rating of product, and Unix time of review. A sample of the electronic product review rating dataset is given below.

Table I: Sample Amazon Product Rating Dataset

Reviewer ID	Product ID	Rating	Timestamp
AKM1MP6P0OYPR	1.33E+08	5	1.37E+09
A2CX7LUOHB2NDG	3.22E+08	5	1.34E+09
A2NWSAGRHCP8N5	4.4E+08	1	1.37E+09
A2WNBOD3WNDNKT	4.4E+08	3	1.37E+09
A1GI0U4ZRJA8WN	4.4E+08	1	1.33E+09

This dataset had 1048576 product reviews for about 1000 products. From this dataset, five products were identified for the analysis presented here. The following proposed algorithm was used to analyze the genuineness of the rating for each product.

Algorithm 1: Computing Average Rating over time

Step I: *Sorting of the timestamps in ascending order.*

Step II: *Determination of the inter review period in Unix time by calculating the difference between the time periods of two consecutive ratings.*

Step III*: Conversion of the time period of each rating into seconds, minutes and then days, thereby attaining the time interval between each rating in days.*

Step IV: *Calculation of the average of all the ratings for each time-period.*

Step V: *Plotting of a graph with the information deduced, with average ratings for each of the time intervals.*

Step VI: *Calculation of the average of all ratings for a product.*

The graphs presented in this section indicate the value of the rating over time. The review rating data of 5 electronic products from the Amazon dataset were used to carry out the analysis presented here. The analysis was done using Excel and Figures 1-5 depict the outcome of the experimental analysis. Table 2 presents a consolidated summary of the average ratings obtains as well as the time span of the complete set of reviews for each of the products.

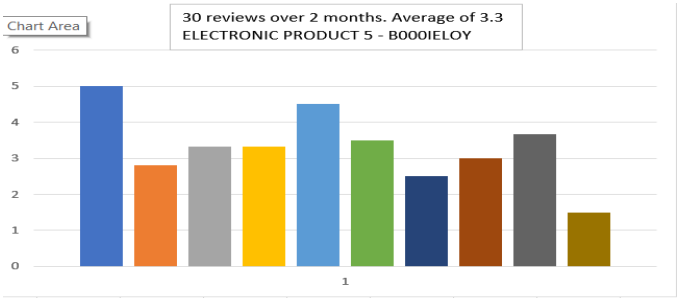


Fig. 2: Average Rating for Product 5

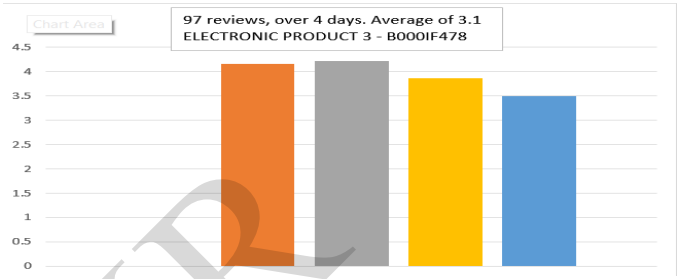


Fig. 3: Average Rating for Product 3

Table II: Analysis of Sample Products

Product Analysis Sample	No. of Reviews	Time Period	Average Review (max. 5)
1	97	4 days	3.1
2	30	2 months	3.3
3	47	1.5 years	4.3
4	66	3 years	4.01
5	156	6 years	4.01

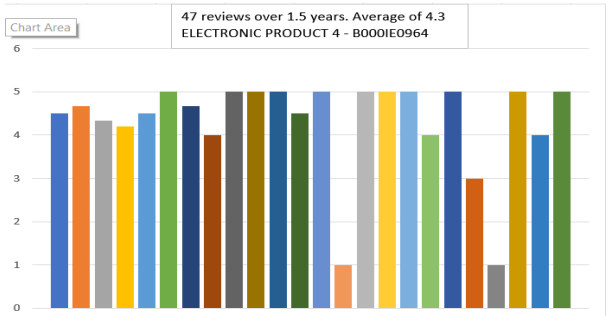


Fig. 4: Average Rating for Product 4

IJBAMR

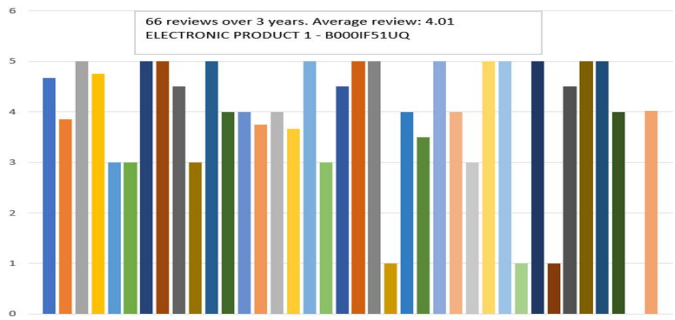


Fig. 5: Average Rating for Product 1

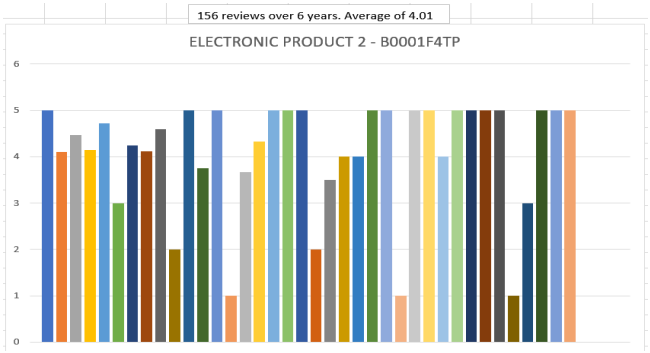


Fig. 6: Average Rating for Product 2

V. CONCLUSION AND FURTHER ENHANCEMENTS

With a growing trend in online shopping there is a need currently to provide an authentic environment for analyzing product reviews and ratings. This research paper expounds a methodology to identify fake ratings among genuine ones over time. This can be used to implicitly identify fake reviewers in cyberspace. Amazon Product Datasets were used for the analysis presented in this paper. An algorithm was proposed for the evaluation of the average rating over time.

A simple classification tool to identify the product specific point on the timeline has also been used in the research work presented in this paper. The proposed fake rating filter therefore helps to tag possible fake reviewers, besides indicating the optimal period for the rating assessment for each product.

REFERENCES

[1] Jure Leskovec and Andrej Krevl, "SNAP Datasets: Stanford Large Network Dataset Collection", June 2014 http://snap.stanford.edu/data

[2] Julian McAuley, University of California San Diego, Amazon Product Data, http://jmcauley.ucsd.edu/data/amazon/

[3] Arjun Mukherjee and Bing Liu, Natalie Glance, "Spotting Fake Reviewer Groups in Consumer Reviews", In proceedings of WWW 2012, April 16–20, 2012, Lyon, France

[4] Arjun Mukherjee, Vivek Venkataraman, Bing Liu and Natalie Glance, "Fake Review Detection: Classification and Analysis of Real and Pseudo Reviews", Technical Report, Department of Computer Science (UIC-CS-2013-03), University of Illinois at Chicago.

[5] Andreas Munzel, "Assisting consumers in detecting fake reviews: The role of identity information disclosure and consensus", Journal of Retailing and Consumer Services, September 2016, Volume 32, pp 96-108.

[6] Feng Zhu, Xiaoquan (Michael) Zhang (2010) Impact of Online Consumer Reviews on Sales: The Moderating Role of Product and Consumer Characteristics. Journal of Marketing: March 2010, Vol. 74, No. 2, pp. 133-148.

IJBAMR