Identifying fake reviews - A heuristics based approach

Anusha Ramamurthy
School of Informatics and Computing
Indiana University Bloomington
anuramam@iu.edu

Abstract

Increased accessibility to Internet and social media platforms has given rise to bots and automated scripts being used to generate content that emulate human behavior. This leads to widespread misinformation about products. Fraudulent reviews result in overnight popularity of some products that were completely unheard of and are possibly of substandard quality, resulting in consumers being unhappy once they receive them. Many online shopping companies lose credibility. Hence, it is important for large companies like Amazon to identify such threats to their business using sophisticated and innovative techniques to combat fraudulent reviews. This project proposes to study the past and current algorithms proposed to identify fraud and suggest areas for improvement.

1 Introduction

Online shopping has reached its peak since 1994. According to Forrester: US online retail sales will reach more than $500 billion by 2020, up from $373 billion in 2016(0). Almost everyday there is a retail company that closes its stores to completely focus on online sales. Companies like Aeropostale have been forced into bankruptcy(0). Most of us shop at Amazon for everything. From books to groceries, we rely on these online retailers for quality goods. More accurately we trust the reviews provided by the previous buyers. The more favorable rating a product has the more likely we are to buy it. However, spammers and fraudsters contaminate this area. With the rise in popularity of Amazon as the go-to website for shopping, there has been a significant rise in new companies that pay individuals to review products. Such incentives create a biased collection of reviews.

2 Related Works

Amazon has dedicated its efforts to stop the rise of fake and incentives reviews by suing companies that pay individuals for writing review's about their products (0). While Amazon may have mechanisms to identify fake reviews it is unclear if or what those measures may be. The ”verified buyer” tag provided by Amazon, helps potential buyers to make a decision to read the review. Another measure is the Amazon Vine Program(0).
Widespread fraudulent news is familiar to all of us. In the last year, during US elections, we have seen a phenomenon of false new stories surfacing social media platforms. There have many attempts and heuristics provided by research scholars to combat fraud in social media platforms. Here are a few of them that are most relevant to the research questions explored in this paper.

- Fake It Till You Make It: Reputation, Competition, and Yelp Review Fraud(0): While this paper analyzes the economic incentives of committing review fraud on Yelp, it sheds light on some interesting patterns about the occurrence of fraud, that helped to analyze the current data set as well. The paper points out a valid point that many reviews may not be written by actual buyers. It also mentions that a restaurant is more likely to hire people to write negative reviews about a popular restaurant, in contrast to the idea that it may hire people to write positive reviews about itself. It also points out that most restaurants hire people to write fake reviews when their own business is going down, and there is a lot of competition.

- Strangers Intrusion Detection - Detecting

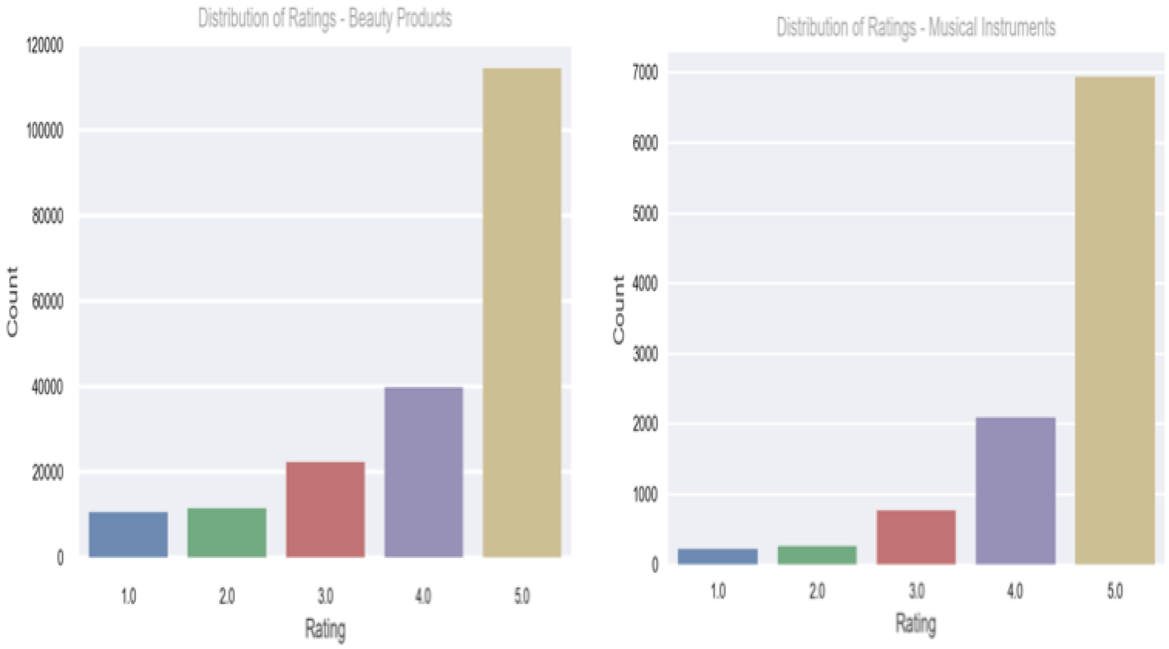

Figure 1: Ratings of Musical Instruments

Spammers and Fake Profiles in Social Networks Based on Topology Anomalies(0): Certain traits can be borrowed from one social media to the other. While Amazon reviews are not directly comparable to Google+ or Academia.edu pages, social behavior on these online platforms may have similar patterns. This paper provides insights on how to identify a fake profile on social media platforms by looking at the topology of social networks.

- Identifying Fake Profiles In LinkedIn(0): This paper sheds light on how to use the metadata associated with profiles to identify fraud. Every review is connected with some metadata about the reviewer and using features similar to those advocated in this paper may help in narrowing down the profiles that are frequently associated with fake reviews across categories. The paper manually tagged accounts considering various aspects like duplicate pictures, similar job descriptions. This is useful in the current scenario where it is useful to identify reviews written by different reviewers but have the same review or style of writing.

- Text-Based User-kNN: Measuring User Similarity Based on Text Reviews(0): This paper reflects upon the use of kNN algorithm to identify similarities in Reviews to create a better recommendation of products. This is relevant to the current research question of identifying reviewers who write fake reviews under different names but their style of writing may inherently be the same.

- Identifying fake Amazon reviews as learning from crowds(0): The authors of this paper highlight the problem that while we may be able to classify a review as fake or not, there is no certainty. There is no mechanism to check the results against an already identified and validated fraudulent review data set. For this reason, the authors of this paper have created a collection of deceptive Amazon book reviews in collaboration with crime writer Jeremy Duns.

3 Data

The Amazon product reviews data set was created by Julian McAuley and the smaller sets are freely available for research(0). This data set contains product reviews and metadata from Amazon,

good
easy
durable
stuff
best
guitar
sturdy
awesome
much
time
great
strings
beginner
bought
works
value
enough
cable
got
buy
just
ever
fits
made
recommend
strap
product
fav
stand
problems
cases
ends
excellent
place
spider
uncomfortable
strong
100
seems
drum
grasp
want
wimpy
years
designed
loved
issues
completely
connected
mics
toys
perfect
paper
nice
plugs
later
weight
day
problem
light
rice
tangle
using
price
home
happy
ear
two
les
beyond
gift
short
quite
sounds
abuse
ordered
know
pedals
provides
job
far
paul
quality
deal
able
picks
like
music
holding
bit
need
lot
cloth
tool
work
thank
towel
explore
bass
price
microphone
stay
protection
quiet
nuff
expected
martin
still
comfortable
pleased
friend
today
little
fine
months
wanted
affordable
useful
think
time

Figure 2: Word Cloud of short length reviews

including 142.8 million review's spanning May 1996 - July 2014. For this project using the smaller 5-core data sets appears to be a good fit as , it is a subset of the data in which all users, and items have at least 5 reviews (41.13 million reviews).

3.1 Statistics

The data set is divided into different product categories. For the purpose of this study, I use the Musical Instruments and Beauty products category. The Musical Instruments Data Set had 10,261 reviews, of which 7 reviews had no characters in the review text column. The reviews have a rating between 1 to 5, with mean 4.5 and standard deviation of 0.89. The Beauty products category data set contains 198,502 reviews of which 27 had no characters in the review text column, with a rating between 1 and 5, with mean rating between 4.19 and standard deviation of 1.19. We see that between the two categories, the ratings are deviate more in the case of Beauty products. This may be because the total number of products under beauty are 12,101 and 900 in the case of Musical Instruments.

3.2 Data Description

Every review in the data set has 10 attributes that provide details that can be used to answer research questions. They are summarized in the below table.

Attribute Name	Description
reviewerID	ID of the reviewer
asin	ID of the product
reviewerName	name of the reviewer
helpful	helpfulness of the review
reviewText	text of the review
overall	rating of the product
summary	summary of the review
unixReviewTime	time of the review
reviewTime	time of the review(raw)

Table 1: Data Description.

4 Methods

Given that the data set contains no filtering or labels to denote a fake review as identified by Amazon, after reading many related studies, this paper proposes heuristics that can be used to label reviews for further moderation and analysis. Thus the results of this paper fall under the category of unsupervised learning.

4.1 Flagging reviews for moderation

4.1.1 Heuristic A - Large Number of Reviews by a Reviewer in a Single Day

Exploring the data set, we observe that a number of reviewers have written reviews on the same day for a variety of products. The rationale behind suspecting such a reviewer as a fraudster is that given general buying patterns, how many of us are likely to buy 5 or higher number of products in the same category? Even if one did buy 5 complimentary products such as say a Guitar, a pick or a tuner and so on the same day, how likely is he or she to submit reviews for all these products on the same day? Since the data set provides no information about the actual date of purchase, we cant cross check. However the data set contains as high as 14 products that were reviewed by a single reviewer on the same day. We flag such reviews for further investigation.

4.1.2 Heuristic B - Short Reviews with less product details

Splitting the reviews into words and removing stop words is used to identify the average function words used in reviews. I was able to identify a number of reviews that have 6 or less function words. A word cloud generated from these reviews shows that good, recommend, works are some of the most repeated words and contain no words about the products themselves.This can be used as one method to flag reviews for moderation. There are certain reviewers who appear exactly once in the entire data set. These one time reviewers write 5 and above reviews in one day and never show up again. I flag them since its entirely possible that it is a contract job, and these reviewers are paid to create new id's, write reviews and never use that id again.

4.1.3 Heuristic C - Reviewers with No Name details

I also flag reviewers who have null fields in place of reviewerName. In the Musical Instruments category, we see 14 reviewers who haven't included Name as part of their account details and 1082 users in the Beauty products category.

4.2 Heuristic D

Looking at the reviewers who have written exactly one review each day, shows highly informational reviews. This appears to be a fraudster or any fake operative to fly under the radar.

4.3 Machine Learning Methods

Once we flag these reviews I have used kNN to identify and flag further such reviews. Since some of the reviews are already part of the training, it is highly possible that kNN will pick up the same reviews again. Hence I sample the data frame with a random state and only take about 40-60% of the reviews, as the data set is quite large.

5 Evaluation

For evaluation of our heuristics, I measure the percentage of such observations in the current data set. The values reported in the table are the percentage of results with respect to the entire population.

Measure	%	Avg.Rating
Blank reviews	0.07	
Missing Names	0.98	
Short Reviews	0.61	4.70
One Time Reviewers	1.38	4.70
One Review a Day	7.78	4.72
Repeated Reviews	0.10	4.98

Table 2: Musical Instruments Category.

Measure	%	Avg.Rating
Blank reviews	0.01	
Missing Names	4.84	
Short Reviews	0.64	4.19
One Time Reviewers	0.95	4.45
One Review a Day	6.37	5.00
Repeated Reviews	0.01	5.00

Table 3: Beauty Products Category.

6 Discussion and Conclusions

6.1 Summary

As shown in this paper, there are significantly high numbers of reviews that are suspect. The ratings all lean towards a perfect score of 5, which is definitely suspect, as given the number of products that are launched nearly every day by new companies, I would have assumed that some of these products would rate lower.

6.2 Limitations

Given the fact that Amazon has not yet let public know the methods they use to identify fake reviews and neither have they released a labeled set of reviews, the heuristics set forth in this paper are no way confirmed to be true. The data set definitely provokes the question or thought that fake or biased reviews exist in this data set. If I had more time, I would like to explore machine learning methods that are able to detect stylistic patterns among the reviewers. This is believe will help validate the heuristics.

6.3 Future Work

I would like to explore using the current methods to flag and moderate reviews in other categories as well. Especially the Books Category has false reviews that were written by authors themselves and confirmed by many, it would be an interesting category to test the heuristics on. It would also be interesting to use an unsupervised Machine Learning Algorithm like k-Means to identify a natural clustering among the reviews, if it exists. It would also help to create newer and better heuristics.

7 Acknowledgments

I would like to thank Dr. Allen Riddell for giving me the opportunity to explore this data set and helping me identify the research avenues to explore.

References

Forrester. Report.

Time newsletter:http://time.com/money/4386499/retail-stores-closing-locations/.

CS Monitor:.

Amazon vine program.

Georgios Zervas Michael Luca. Fake it till you make it: Reputation, competition, and yelp review fraud, michael luca, georgios zervas. *Management Science*, July 11, 2015.

Yuval Elovici Michael Fire, Gilad Katz. Strangers intrusion detection - detecting spammers and fake profiles in social networks based on topology anomalies.

Kaushik Dutta Shalinda Adikari. Identifying fake profiles in linkedin.

Ferrario MA. Whittle J. Terzi M., Rowe M. Text-based user-knn: Measuring user similarity based on text reviews. Springer, Cham, (2014).

Massimo Poesio Tommaso Fornaciari. Identifying fake amazon reviews as learning from crowds.

J. Leskovec J. McAuley, R. Pandey. Inferring networks of substitutable and complementary products. Knowledge Discovery and Data Mining,, 2015.