Detecting Opinion Spammer Groups

Lars Stelzer
lars.stelzer@studium.uni-hamburg.de
7346178

Niklas von Boguszewski
nvboguszewski@googlemail.com
6790872

Knowledge Processing in Intelligent Systems: Practical Seminar
Knowledge Technology, WTM, Department of Informatics, University of Hamburg

Abstract—This paper aims to detect opinion spammer groups on the Yelp dataset.
Methode nennen
results zeigen (conclusion)

I. Introduction

Consumers' decisions are significantly influenced by product reviews. Due to the ongoing digitalization more and more consumers rely their decision making on online reviews. Online product reviews have by now become the second must trusted source of product information, followed by recommendations from family and friends. This happens due to the belief that online product reviews reflect recommendations of "real" people. A rising online marketing budget has led to a larger influence and importance of online reviews in purchase and decision making. By the fact that positive/negative reviews are able to enhance/defame products organised opinion spamming started to take place. Nowadays it is estimated that up to 20% of online reviews could be fake.

We use Yelp as the basis of this papers approach since it is one of the largest online consumer review hosting sites for services (e.g., restaurants, hotels, etc.) in the commerical setting.

Due to the fact that we are simply not able to declare a user or a spammer group as definite "spam" we are trying to identify the groups that are suspicious and most likely to be spammers.

The temporal dynamics in which spammer groups operate is unclear.This chapter shows the motivation of why this topic is important.

Wir haben user mit unserer methode gefunden, die highly suspicous sind und dann ein beispiel nennen

warum sind fake reviews doof

was ist der impact

leute treffen wirtschaftliche entscheidungen vermehrt anhand der reviews

reviews tragen zur entscheidungsfindung bei und dabei ist ein verlässliches raiting essentiell

viele fake reviews zersören das vertrauen in Yelp und ähnliche Plattformen wie bspw. Amazon.

yelp hat 20 prozent fake reviews und filtert selber

es sind noch fake reviews da

weitere fake reviews reduzieren

gruppenansatz von paper CPM XY Twist nennen

story telling eine suspicous gruppe nennen (wo jeder sagen würde es ist klar spam) dann glaubt uns jeder

idee warum gruppen –> Venn Diagramm zeigen

II. Related Work

Over the last years the field of spam review detection has been experienced plenty of attention. The existing researches can be divided based on the data and features they analyze. Below are presented three different approaches to filter and detect spam.

1. Review spam detection: this approach is based on the data of the review (e.g., textual content, meta data, etc.) to detect and filter spam reviews.
2. Review spammer detection: this approach is aiming to find the spammer it self. This can be, for example, achieved by identifying atypical behavior, such as posting time or geolocation.
3. Group spammer detection: goal with this approach is to identify whole group of spammers due to the fact that group spammer are more influential and harmful than individual spammer.

Was haben andere gemacht

single fake review erkenne (3)

gruppen fake reviews erkennen

ergebnisse ...

gruppenansatz von paper CPM XY 2 andere gruppenansätze (man kann Twist nennen)

—- In this section we show important work which has been made by other researchers like [2] and [1].

In this section we also want to highlight why this approach can be beneficial for this research area.

III. Yelps Review System

In this section we shortly describe how Yelps review system works.

you can give stars to business

free for all

yelp filters reviews on its own

freitext

yelp hat 20 fake reviews
yelp filterd selber

IV. Dataset

The dataset we use is publically available on Yelp.com and consists of real-world reviews about different businesses (e.g., restaurants, hotels, etc.). Yelp is a recommendations platform based on user-generated reviews. The dataset contains information about reviews, businesses, user data, and geographical data. More than 8 million reviews and more than 200.000 businesses are included in the dataset. For this paper, we look at reviews of 2016 that appeared in the area of Charlotte in North Carolina USA. In total, we examine 51261 reviews and 5954 businesses.

V. Feature

BST
t=28
Review Count for each user
Cosine Similarity
Duplicates/Near Duplicates (beta>0.7)
In this section we descripe approximately 3 features we want to experiment with.
Extreme Rating ( ):

VI. Model

In this section we describe our model/graph which aims to find opinion spamming groups.
1. undirected Graph (G) 2. We look at the reviews of each business 3. Then, if 2 reviewer commented on the same business within in specific time window (alpha=6, in days) and the same star rating, then we add them to the Graph G 4.
Twist nennen (+ Grenzwertvariablen)
Image of Groups

VII. Results/Analysis

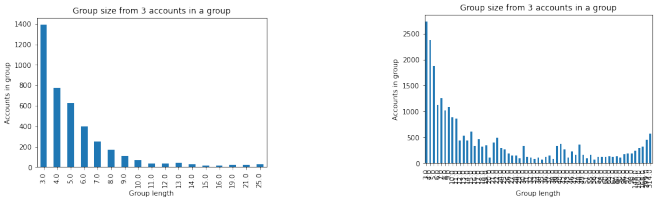
In this section we show our plots and describe the results/interpretation.
look at artecicial groups
- wie gut finden wir unseren suspicious score - reasoning über correlation - reasoning über vermutlichen prozentualen anteil -

VIII. Conclusion

Twist nennen und sagen ob es eine gute Idee war
Finale interpretations der Ergebnisse - Was haben wir eigentlich gemacht - was war das ziel - gruppen erkennen, die fake reviews schreiben - haben wir das erreicht
Was machen wir nächstes mal anders + Empfehlungen

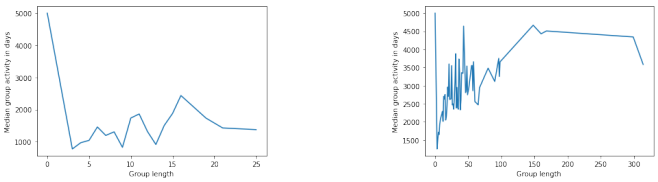
IX. Experimental plots

200.000 rows of the review json



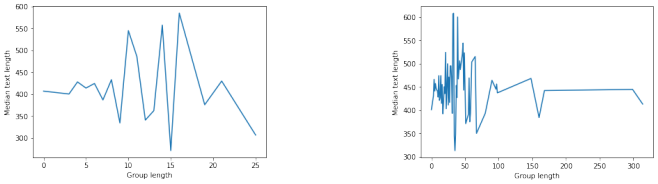
(a) Positive (b) Negative

Figure 1: Groups



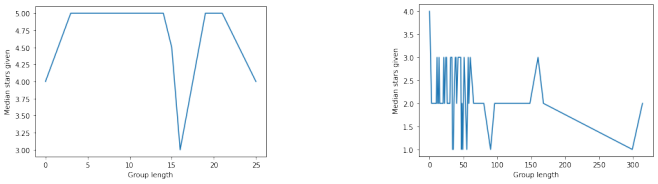
(a) Positive (b) Negative

Figure 2: Activity



(a) Positive (b) Negative

Figure 3: Length



(a) Positive (b) Negative

Figure 4: Stars

References

[1] Euijin Choo, Ting Yu, and Min Chi. Detecting opinion spammer groups through community discovery and sentiment analysis. In IFIP annual conference on data and applications security and privacy, pages 170–187. Springer, 2015.

[2] Arjun Mukherjee, Abhinav Kumar, Bing Liu, Junhui Wang, Meichun Hsu, Malu Castellanos, and Riddhiman Ghosh. Spotting opinion spammers using behavioral footprints. In Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining, pages 632–640, 2013.