Network-Based Spam Detection Framework For Reviews And Credit Card Fraudulent

D.J.Hani Mary Sheniha[1], N.Gokulnath[2], A.Harish kumar[3]
[1, 2, 3] Dept of Computer Science
[1, 2, 3] JEPPIAAR SRR Engineering College, Chennai

Abstract- *Social media plays a major role in user decisions using reviews, feedbacks. The possibility that anybody can leave a review provide a golden opportunity for spammers to write spam reviews about products and services for different interests. Finding out these spammers is a main content of research and although a valued number of studies have been done recently according this end, but so far the methodologies put forth still barely detect spam reviews, and none of them show the importance of each extracted feature type. Every spam detector system fails to identify this kind of spammers or at least has some trouble to spot them. Two categories effectively named review-behavioural and user-behavioural to identify spam reviewers and credit card fraud intruders. Using web services available to build this application and accuracy in detecting spam reviewers and credit card fraud intruders.*

Keywords- Social Media, Social Network, Spammer, Spam Review, Fake Review, credit card fraud intruders.

I. INTRODUCTION

Social Media play an vital role in information propagation which is considered as an important source for producers in their advertising campaigns as well as for customers in selecting products and services. In recent years, people rely on the written reviews in their decision-making processes, and positive/negative reviews them in their selection of products and services. A considerable amount of literature has been published on the methods used to identify spam and spammers. From[1],These methods can be classified into different categories; some using linguistic patterns in text which are mostly based on bigram, and others are based on behavioral patterns that rely on features extracted from patterns in users' behavior which are mostly metadata based and even some methods using graphs and classifiers .Despite this great deal of efforts, many aspects have been missed or remained unsolved. One of them is a classifier that can calculate feature weights that show each feature's level of importance in determining spam reviews. The general concept of the proposed framework is to model a given review dataset as a Heterogeneous Information Network and to map the problem of spam detection into a classification problem. Dataset as in which reviews are connected through different node types. A weighting algorithm is then employed to calculate each feature's importance. These weights are utilized to calculate the final labels for reviews using both unsupervised and supervised approaches. Cash and cheque use is dropping as consumers continue to choose digital payments. Users are embracing the convenience of online shopping continues to grow and mobile payments are gaining traction. Globally, online sales continue to account for a growing portion of total retail. By 2020, forecasts estimate that online shopping will have doubled since 2015 and will account for 14.6% of all retail sales. This trend is clearly reflected, where online spending is growing five times faster than traditional retail spending. In line with global trends, the shift towards online shopping has also seen a rise in online payment fraud. Card-not-present fraud accounted of all fraud on transaction cards. Notwithstanding this rise in online fraud, action taken by industry to combat fraud is taking effect in many countries [2]. This is reflected in the growth rate of CNP fraud, which has decreased over the last two years.

II. EXISTING SYSTEM

Spam detection is a relatively new research problem. A preliminary study was reported. In this work, three types of review spam were identified, namely untruth-full reviews, re-views on brands and non-reviews. In some of the specified, negative review detection is performed by using fake reviews. Reviews specially come up with ratings and special characters. Detecting fraudulent ratings has been in several works including. The techniques used are clustering ratings into unfairly high ratings and unfairly low ratings, and using third party ratings on the producers of ratings and ratings from less reputable producers are then deemed as un-fair. Ratings are found, they can be deleted to restore a fair item in system. These works do not address spammer detection directly. They will not conduct evaluation of their techniques on original data.

Review helpfulness prediction is closely related to review spam detection. The former aims to differentiate reviews of different helpfulness. A helpful review is one that

is informative and useful to the readers. The purpose of predicting review helpfulness is to help review sites to give feedbacks to the review contributors and to help readers choose and read high quality reviews. A classification approach to solving helpfulness prediction using review content and meta-data features was developed. The meta-data features use dare review's rating and the difference between the review rating and the average rating of all reviews of the product. It proposes to derive from reviews content features that correspond to informativeness, readability and subjectiveness aspects of review. These features are then used to train a helpfulness classification method. All e-commerce companies suffer and face legal issues due to increase of credit card frauds. Actually 20% users in Yelp are spammers. So far these techniques barely detect spam reviews, and none of them show the importance of each extracted feature type. The fact that anyone with any identity can leave comments as review provides a tempting opportunity for spammers to write fake reviews designed to mislead users' opinion.

Along with the great increase in credit card transactions, credit card fraud has been increased. The fraud is one of the major causes of great financial losses, not only former chants, individual clients are also affected. Three methods to detect fraud are presented. Gaussian mixture model is used to model the probability density of credit card user's past behavior so that the probability of current behavior can be calculated to detect any abnormalities from the past behavior. It is used to describe the statistics of a user and the statistics of different fraud. Building an accurate and easy handling credit card risk monitoring system is one of the key tasks for the merchant banks. The models are compared in terms of their performances. To improve the fraud detection system, the combination of the three presented methods could be beneficial. It is likely to Bayesian Networks based on the input representation method and the developed clustering model in the detection system.

III. PROPOSED SYSTEM

Detection and prevention of spam reviewers and credit card fraud intruders is done efficiently. Using web services available the application is developed and provides high accuracy in detecting spam reviewers and credit card fraud intruders [3]. Analyze reviewer behaviors and reviewers history and review written on the same released date of the product /item/service. As review with same number of exclamations and percentage of capital words used in the review and usually spammers hide their identity for security reasons.

In proposed system, parallel granular neural networks is used for developing the rapidity of data mining and information detection process in credit card fraud detection, From[4]. A complete system has been implemented for this purpose. It also suggest a credit card fraud detection system (FDS) using meta learning techniques to learn models of fraudulent credit card transactions. General strategy that provides a means for combining and integrating a number of separately built classifiers or models.Automated validation of user e-mail address, user billing address, user phone number, user IP address and user device ID are the technique used [5].Manually checking user integrity by checking social media profile and validating user business website, architecture diagram shown in fig 1.Increase of fraud which results in loss each year, several modern techniques in detecting fraud are persistently evolved and applied to many business fields. Fraud detection involves monitoring the activities of populations of users in order to estimate, perceive or avoid undesirable behavior. Unwanted behavior is a wide term including criminal behavior, fraud, Intrusion, and account defaulting. It presents a survey of current techniques used in credit card fraud detection. The goal is to provide a comprehensive review of different techniques to detect fraud.

IV. ARCHITECTURE DIAGRAM

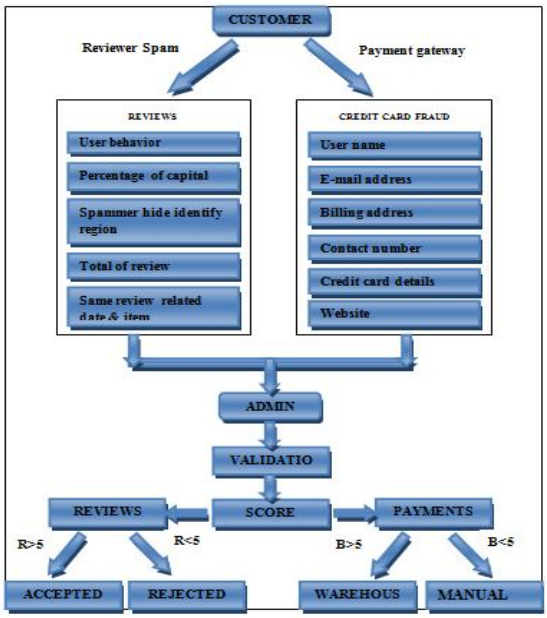

Fig 1: Architecture diagram for netspam.

IV. MODULES

1. User authentication.
2. Review Spammers.
3. Payment gateway.

4. Automated obtaining and validating user.

1. User authentication

Each user behavior is been predicted. User past activities are been analyzed. If the user is very first time giving the review, the user is more analyzed against genuine user or spammer. In this module, user name (first and last name) are been validated. The user name is checked against Face book, LinkedIn and other Google forums whether the user is active checking their photos, friends, timelines etc. User requires fast algorithms that predict validity of email address. Such methods depend heavily on heuristic algorithms and statistical models as shown in fig 2.

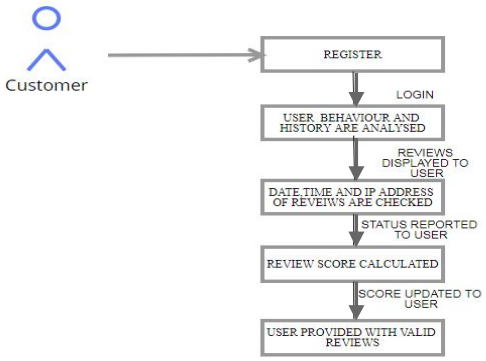

Fig 2: User Authentication

***2.* Review Spammers.**

In this module, each user behavior is been predicted. User past activities are 'been analyzed. If the user is very first time giving the review, the user is more analyzed against genuine user or spammer. Reviews with same released date related to Item. This is defined as Review-Early Time Frame-Review. In this module whenever product and review is been added, the date is been analyzed and kept tracked. Spammers put '!' in their sentences as much as they can to increase impression on users and highlight their reviews among other ones. So many spammers use capital letters to obtain the spotlight. Spammers often write spam reviews hiding their true identity which is known as camouflage. For example they write positive reviews for restaurant or negative reviews for low-quality ones. Every spam detector system fails to identify this kind of spammers or at least has some trouble to spot them as shown in fig 3.

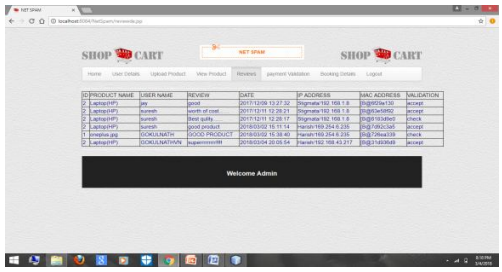

Fig 3: Review spammers.

3. Payment gateway.

In this module, customer e-mail address is been checked and validated. Email addresses are often requested as input to website as user identification for the purpose of data validation. Email validations are given based on the API endpoint. Large websites, bulk mailers and spammers require fast algorithms that predict validity of email address. Such methods depend heavily on heuristic algorithms. The Address Verification System (AVS) is a system used to verify the address of a person claiming to own a credit card. The system checks the billing address of the credit card provided by the user and address on file at the credit card. AVS verifies that the billing address of the credit or debit card matches the address that was given by the customer. In this project, verified billing Places using library in the Google Maps JavaScript API as shown in fig 4.

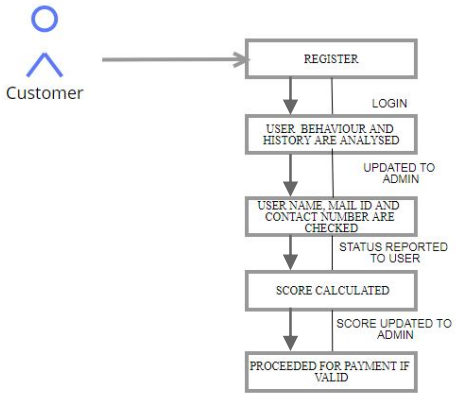

Fig 4: Payment gateway.

4. Automated obtaining and validating user.

Phone number validation helps you reach customers and validate them. The powerful tool utilizes data from over 200 telecommunications providers to deliver accurate, fast results using web services. It even cleanse phone numbers at point-of-entry, so you can be sure they're valid before they enter your database. Phone validation tool ensures that accurate telephone numbers are collected to keep this valuable channel open. It also detect the location and line type to

ensure. An IP address consists of four numbers separated by periods. The configure of an IP address is a 32-bit numeric address written as four decimal numbers separated by periods; each number can be written as 0 to 255.A MAC address is a unique identifier assigned to most network adapters or network interface cards by the manufacturer for identification.This format gives 281,474,976,710,656 possible unique MAC addresses. In this module, user name are been validated. The user name is checked against Facebook, LinkedIn and other Google forums whether the user is active checking their photos, friends, timelines as shown in fig 5.

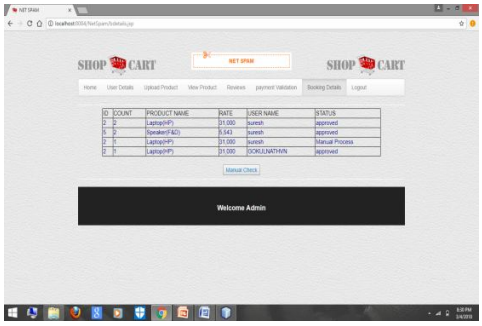

Fig 5: Validating the payment.

V. LIMITATION OF EXISTING SYSTEM

- Spammer reviews can potentially bang credibility and affect economic losses.
- Written negative reviews intended to mislead users' opinion.
- A crawler tool needed which is costly and no open source.
- Accuracy of the spam detector is low.

VI. APPLICATIONS

- Payment related websites.
- E-commerce and M-commerce application.
- Used in review based websites

VII. CONCLUSION

The proposed system provides very effective in identifying spam reviews and leads to a better performance. In addition, It also detect and prevent credit card fraudulent as well. This proposed system provides efficient and accurate architecture for detecting and preventing internet spammers. Detection and prevention of spam reviewers and credit card fraud intruders is done efficiently. Methods are effectively named review-behavioral and user-behavioral to identify spam reviewers and credit card fraud intruders. The score analyzing technique is performed to identify the spammers by phone call verification and by this identification the order is sent to the warehouse for processing and shipping.

VIII. FUTURE ENHANCEMENTS

For future work, meta path concept can be applied to other problems in this field. For example, similar framework can be used to find spammer communities. For finding community, reviews can be connected through group spammer features and reviews with highest similarity based on meta path concept are known as communities. In accumulation, utilize the manufacturing features is an interesting future work on this study as used features more related to spot spammers and spam reviews. Moreover, while single networks have received considerable attention from various disciplines for over a decade, information diffusion and content sharing in multilayer networks is still a young Research. When entering payments from Debit Card receipts the only option right now is "Check" which automatically fills in next check #. It would be very time effective to have option for new custom type.

REFERENCES

[1] Saeedreza Shehnepoor, Mostafa Salehi, Reza Farahbakhsh, Noel Crespi, "NetSpam: a Network-based Spam Detection Framework for Reviews in Online Social Media", IEEE Transactions on Information Forensics and Security, Volume:12, Issue:7, pp. 1585-1595, July 2017.

[2] Renu, Suman, "Analysis on Credit Card Fraud Detection Methods", International Journal of Computer Trends and Technology, Volume: 8, Issue: 1, Feb 2014.

[3] Khyati Chaudhary, Bhawna Mallick, "Credit Card Fraud: The study of its Impact and detection techniques", International Journal of Computer Science and Network, Volume: 1, Issue: 4, August 2012.

[4] Khyati Chaudhary, Jyoti Yadav,Bhawna Mallick, "A review of Fraud Detection Techniques: Credit Card", International Journal of Computer Applications ,Volume: 45, Issue: 1, pp. 0975 – 8887, May 2012 .

[5] W.A. Awad, S.M. ELseuofi, "Machine Learning Methods For Spam E-Mail classification", International Journal of Computer Science & Information Technology, Volume: 3, Issue: 1, Feb 2011.

[6] V. Dheepa, Dr. R. Dhanapal, "Analysis of Credit Card Fraud Detection Methods", International Journal of Recent Trends in Engineering, Volume: 2, Issue:3, November 2009.

[7] Abhinav Srivastava, Amlan Kundu, Shamik Sural, Arun K. Majumdar, "Credit Card Fraud Detection Using Hidden Markov Model", IEEE Transactions on Dependable and Secure Computing, vol. 5, no. 1, January-March 2008.

[8] L. Mukhanov, “Using bayesian belief networks for credit card fraud detection,” in Proc. of the IASTED International conference on Artificial Intelligence and Applications, Insbruck, Austria, pp. 221– 225, Feb. 2008.
[9] Abhinav Srivastava, Amlan Kundu, Shamik Sural and Arun K. Majumdar, "Credit Card Fraud Detection Using Hidden Markov Model", IEEE Transactions On Dependable And Secure Computing, Volume: 5, Issue: 1 , January-March 2008.
[10] Hao Zhang, Alexander C. Berg, Michael Maire, and Jitendra Malic. "SVM-KNN: Discriminative nearest neighbour classification for visual category recognition", IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2006.
[11] Vladimir Zaslavsky and Anna Strizhak, “credit card fraud detection using self organizing maps”, information &security. An International Journal, Volume:18, 2006.