

Opinion Spam Detection, Tools and Techniques : A Review

Hema Dewangan

*Rungta College of Engineering and Technology
Dept. of Computer Science and Engineering
Bhilai, Chhattisgarh, India.*

Prof. Om Prakash Dewangan

*Rungta College of Engineering and Technology
Dept. of Computer Science and Engineering
Bhilai, Chhattisgarh, India.*

Abstract

An excellent source of collecting the reviews on specific product is various online shopping sites where people share their reviews on products and their shopping experience. People may come through the wrong opinions known as review spam. Therefore, for this it is essential to detect it by some means. In this paper, a survey is done on the various techniques introduced to detect the spam of user review. This paper also compares various tools and methods available for detecting spams and spammers.

Keywords: Review spam, un-truthful reviews, opinion spam, rating spam.

I. INTRODUCTION

At the present, there is no quality control for social networking sites and one has having freedom to share their reviews on social networking sites which helps to lead the review spam. And it is a requirement to recognize review spam because most of the users make their decision based on the reviews. This condition mainly arises for various online shopping sites or the sites or hotels also. Various techniques are introduced and used for detecting review spam.

There are mainly 3 types of review spam, which are

Type 1: Un-truthful review or false opinions which is divided mainly into two category i.e., Positive Spam review (Undeserving opinion to promote product) and Negative Spam Review (negative opinion to damage reputation)

Type 2: Reviews on brand (reviews on some particular brands)

Type 3: Non-reviews (contain no reviews) which is divided in Advertisements, any Random text, Question or answers or Comment on other reviews.

Many copies of similar content is being distributed among multiple users in the internet. These messages are of no use because of duplication. The importance of message is reduced due to the duplication of many similar messages. Most of the spams are present on the rating websites such as e-commerce, movie rating, product rating etc. People misuse platform and try to capture audience attention by providing many fake reviews about the product.

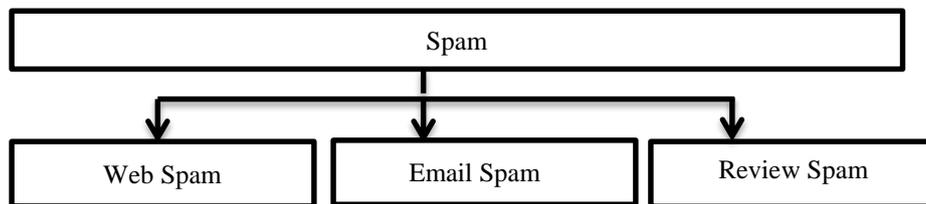


Fig.1. Various types of spam

Web Spam

The main purpose of web spam is to make people follow some untrusted link. These links will steal all the user information and send data to the hacker server [1].

Email Spam

A substitute kind of spam is email spam, which is furthermore not the same as review spam [2]. Email spam (furthermore called rubbish messages) incorporates getting the uninvited business plugs.

Review Spam

If anyone wants to buy some product let say from Flipkart. The user will look for its review [3]. Spam Reviews refers to "illegal" activities that try to mislead readers by giving false negative opinions to some other entities in order to damage their reputations

II. LITERATURE SURVEY

E. Zheleva et. al. [14], Spam is a developing issue; it interferes with legitimate email and weights both email users and service providers. In this work, we propose a receptive spam-sifting framework in view of columnist notoriety for use in conjunction with existing spam-separating procedures. The framework has a trust-support part for clients, in light of their spam-announcing conduct. The test that we consider is that of keeping up a dependable framework, not powerless against noxious clients, that will give the early spam-battle location to diminish the expenses brought about by clients and frameworks.

J. Martinez-Romo et. al. [15], This paper applies a language model way to deal with various wellsprings of data removed from a Web page, keeping in mind the end goal to give great pointers in the identification of Web Spam. Two pages connected by a hyperlink ought to be topically related, despite the fact that this was a feeble logical connection. Hence we have examined distinctive wellsprings of data of a Web page that has a place with the setting of a connection and we have connected Kullback-Leibler disparity on them for portraying the connection between two connected pages.

Nitin Jindal et.al. [16], It is presently a typical practice for web-based business Web locales to empower their clients to compose surveys of items that they have bought. Such audits give profitable wellsprings of data on these items. They are utilized by potential clients to discover feelings of existing clients before choosing to buy an item. They are additionally utilized by item makers to recognize issues of their items and to discover aggressive knowledge data about their rivals.

Gordon V. Cormack [17], Author consider the issue of substance based spam sifting for short instant messages that emerge in three settings: versatile (SMS) correspondence, blog remarks, and email outline data, for example, may be shown by a low bandwidth customer. Short messages frequently comprise of just a couple words and in this way display a test to customary sack of-words based spam channels. Utilizing three corpora of short messages and message fields gotten from genuine SMS, blog, and spam messages, we assess include based and pressure show based spam channels.

G. Cormack [18], Eleven variations of six generally utilized open-source spam channels are tried on a sequential grouping of 49086 email messages gotten by a person from August 2003 through March 2004. Our approach varies from those beforehand revealed in that the test set is expensive, involves uncensored crude messages, and is introduced to each channel successively with incremental criticism. Misclassification rates and Receiver Operating Characteristic Curve estimations are accounted for, with factual certainty interims. Quantitative outcomes show that

substance based channels can take out 98% of spam while bringing about 0.1% genuine email misfortune.

Manuel Egele [19], Web spam indicates the control of pages with the sole plan to bring their position up in internet searches rankings. In this paper, we acquaint an approach with identifying web spam pages in the rundown of results that are returned by a web search tool. In an initial step, we decide the significance of various page elements to the positioning in Web crawler comes about. In view of this data, we build up a grouping strategy that utilizes vital element to effectively recognize spam locales from honest to goodness sections.

Gilad Mishne [20], Author show an approach for identifying join spam normal in blog remarks by looking at the dialect models utilized as a part of the blog entry, the remark, and pages connected by the remarks. Rather than other connection spam separating approaches, our technique requires no preparation, no hard-coded administer sets and no information of finish web availability. Preparatory investigations with ID of regular blog spam indicate promising results.

Lourdes Araujo [21], web spam is a significant issue for web crawlers on the grounds that the nature of their outcomes can be extremely corrupted by the nearness of this sort of page. In this paper, we introduce an effective spam discovery framework in view of a classifier that consolidates new connection based elements with dialect demonstrate (LM)- based ones. These components are not just identified with quantitative information separated from the Web pages, additionally to subjective properties, for the most part of the page joins. The outcome is a framework that altogether enhances the location of Web spam utilizing fewer components, on two expansive and open datasets, for example, WEBSpAM-UK2006 and WEBSpAM-UK2007.

Andras A. Benczur [22], Spammers plan to build the PageRank of certain spam pages by making countless indicating them. We propose a novel technique in light of the idea of customized PageRank that distinguishes pages with an undeserved high PageRank esteem without the need of any sort of white or boycotts or different methods for human mediation. Our strategy is tried on a 31 M page slither of the area with a physically arranged 1000-page stratified arbitrary example with predisposition towards huge PageRank values.

Carlos Castillo [23], Author portray the WEBSpAM-UK2006 gathering, an expansive arrangement of Web pages that have been physically commented on with names showing if the hosts are incorporate Web spam angles or not. This is the principal openly accessible Web spam gathering that incorporates page substance and joins, and that has been marked by an extensive and various arrangements of judges.

III. CONCLUSION

This paper basically studies the problem of singleton review spam detection, which is both challenging and important to solve. This paper shows different approaches for review spam detection. All approach has some advantage as well as some disadvantage. Main intention is to correctly classify the review as a spam or not. This paper also compares various methods and techniques.

REFERENCES

- [1] S. Kumar, X. Gao, I. Welch and M. Mansoori, "A Machine Learning Based Web Spam Filtering Approach," 2016 IEEE 30th International Conference on Advanced Information Networking and Applications (AINA), Crans-Montana, 2016, pp. 973-980.
- [2] P. Rajendran, M. Janaki, S. M. Hemalatha and B. Durkananthini, "Adaptive privacy policy prediction for email spam filtering," 2016 World Conference on Futuristic Trends in Research and Innovation for Social Welfare (Startup Conclave), Coimbatore, 2016, pp. 1-4.
- [3] B. Angsumalee, N. Sotthisopha and P. Vateekul, "A framework to detect unqualified restaurant reviews," 2016 Eighth International Conference on Knowledge and Systems Engineering (KSE), Hanoi, 2016, pp. 115-120.
- [4] S. Xie, Q. Hu, J. Zhang and P. S. Yu, "An effective and economic bi-level approach to ranking and rating spam detection," 2015 IEEE International Conference on Data Science and Advanced Analytics (DSAA), Paris, 2015, pp. 1-10.
- [5] W. Shi, M. Xie and Y. Huang, "Collaborative spam filtering technique based On MIME fingerprints," 2011 9th World Congress on Intelligent Control and Automation, Taipei, 2011, pp. 225-230.
- [6] Nitin Jindal and Bing Liu, "Analyzing and Detecting Review Spam", Seventh IEEE International Conference on Data Mining 2007.
- [7] Snehal Dixit & A.J.Agrawal, "Review Spam Detection", International Journal of Computational Linguistics and Natural Language Processing Vol 2 Issue 6 June 2013 ISSN 2279 -0756.
- [8] Siddu P. Algur, Amit P.Patil, P.S Hiremath, S. Shivashankar, "Conceptual level Similarity Measure based Review Spam Detection", 2010 IEEE.
- [9] C.L. Lai, K.Q. Xu, Raymond Y.K. Lau, Y. li, L. Jing "Toward A Language Modeling Approach for Consumer Review Spam Detection", International Conference on E-Business Engineering 2010.

- [10] Raymond Y. K. Lau, S. Y. Liao, Ron Chiwai Kwok, Kaiquan Xu, Yunqing Xia, Yuefeng Li, “Text Mining and Probabilistic Language Modeling for Online Review Spam Detection”, *ACM Trans. Manag. Inform. Syst.* 2, 4, Article 25, December 2011.
- [11] Sumit Sahu, Bharti Dongre, Rajesh Vadhwani, — Web Spam Detection Using Different Features — in *IJSCE*, ISSN : 2231-2307 , Volume-1, Issue -3, July 2011.
- [12] Ying Liu ,Jian Jin, Ping Ji, Jenny A. Harding, Richard Y.K. —Identifying helpful online reviews: A product designer’s perspective —*Computer-Aided Design* 45 (2013) 180–1940010-4485/\$ –2012 Elsevier .
- [13] C.L. Lai, K.Q. Xu, Raymond Y.K. Lau and Yuefeng Li, —High Order Concept Associations Mining and Inferential Language Modeling for Online Review Spam Detection, 978-0-7695-42577/10 , IEEE, 2010 .
- [14] E. Zheleva, A. Kolcz, and L. Getoor. Trusting spam reporters: A reporter-based reputation system for email filtering. *ACM Transactions on Information Systems*, 27:3.1–3.38, 2008.
- [15] J. Martinez-Romo and L. Araujo. Web spam identification through language model analysis. In *Proceedings of the Fifth International Workshop on Adversarial Information Retrieval on the Web*, pages 21–28, 2009.
- [16] N. Jindal and B. Liu. Review spam detection. In *Proceedings of the 16th International Conference on World Wide Web*, pages 1189–1190, 2007.
- [17] G. Cormack, J. Hidalgo, and E. Sanz. “Spam filtering for short messages”. In *Proceeding of the 16th ACM CIKM*, pages 313– 319. ACM, 2007.
- [18] G. Cormack, J. Hidalgo, and E. Sanz. Online supervised spam filter evaluation. *ACM Transactions on Information Systems*, 25:11.1–11.31, 2007.
- [19] Manuel Egele , Clemens Kolbitsch , Christian Platzer, —Removing web spam links from search engine results in Springer-Verlag France 2009 *J Comput Virol* (2011) 7:51–62.
- [20] Gilad Mishne, David Carmel, and Ronny Lempel. Blocking blog spam with language model disagreement. In *Proceedings of the 1st International Workshop on Adversarial Information Retrieval on the Web (AIRWeb)*, Chiba, Japan, 2005.
- [21] Lourdes Araujo and Juan Martinez-Romo —”Web Spam Detection: New Classification Features Based on Qualified Link Analysis and Language Models” in *IEEE Transactions on Information Forensics and Security*, Vol. 5, no. 3, September 2010.

- [22] András A. Benczúr, Károly Csalogány, Tamás Sarlós, Máté Uher —SpamRank – Fully Automatic Link Spam Detection Work in progress|| in Proc. First Int. Workshop on Adversarial Information Retrieval on the Web (AIRWeb, Chiba, Japan, 2005, pp. 25–38.
- [23] C. Castillo, D. Donato, L. Becchetti, P. Boldi, S. Leonardi, M. Santini, and S. Vigna, —A reference collection for web spam,|| SIGIR Forum, vol. 40, no. 2, pp. 11–24, 2006.

