Quest Journals
Journal of Software Engineering and Simulation
Volume 8 ~ Issue 11 (2022) pp: 16-23
ISSN(Online) :2321-3795 ISSN (Print):2321-3809
www.questjournals.org

Research Paper

Fake Product Review Monitoring System Using Netspam Framework

Prof.Mohammed juned Shaikh Shabbir
Mauli Group of Institution's college of engineering and technology shegaon

Abstract: *Todays, a major part of everyone trusts on content in social media like opinions and feedbacks of a topic or a hotel. The liability that anyone can take off a survey give a brilliant chance to spammers to compose spam surveys about hotels and services for various interests. Recognizing these spammers and the spam content is a wildly debated issue of research and in spite of the fact that an impressive number of studies have been done as of late toward this end, yet so far the procedures set forth still scarcely distinguish spam reviews, and none of them demonstrate the significance of each extracted feature type. In this investigation, propose a novel structure, named NetSpam, which uses spam highlights for demonstrating review datasets as heterogeneous information networks to design spam detection method into a classification issue in such networks. Utilizing the significance of spam features help all to acquire better outcomes regarding different metrics on review datasets. The outcomes demonstrate that NetSpam results the existing methods and among four categories of features; including review-behavioral, user-behavioral, review-linguistic, user-linguistic, the first type of features performs better than the other categories. The contribution work is when user search query it will display all top-k hotels as well as recommendation of the hotel for particular user's point of interest.*

Keywords*-Social Media, Social Network, Spammer, Spam Review, Fake Review, Heterogeneous Information Networks.*

R*eceived 15 Nov., 2022;* ***R****evised 28 Nov., 2022;* ***A****ccepted 30 Nov., 2022 © The author(s) 2022.*
P*ublished with open access at* ***www.questjournals.org***

I. INTRODUCTION

Online Social Media portals play an influential role in information propagation which is considered as an important source for producers in their advertising campaigns as well as for customers in selecting hotels and services. In the past years, people rely a lot on the written reviews in their decision-making processes, and positive/negative reviews encouraging/discouraging them in their selection of hotels and services. In addition, written reviews also help service providers to enhance the quality of their hotels and services. These reviews thus have become an important factor in success of a business while positive reviews can bring benefits for a company, negative reviews can potentially impact credibility and cause economic losses. 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. These misleading reviews are then multiplied by the sharing function of social media and propagation over the web.

1.1 Background

Online Social Media portals play an influential 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 the past years, people rely a lot on the written reviews in their decision-making processes, and positive/negative reviews encouraging/discouraging them in their selection of products and services. In addition, written reviews also help service providers to enhance the quality of their products and services. These reviews thus have become an important factor in success of a business while positive reviews can bring benefits for a company, negative reviews can potentially impact credibility and cause economic losses. 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. These misleading reviews are then multiplied by the sharing function of social media and propagation over the web. The reviews written to change users' perception of how good a product or a service are considered as spam [11], and are often written in exchange for money 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 our proposed framework is to model a given review dataset as a Heterogeneous Information Network (HIN) [19] and to map the problem of spam detection into a HIN classification problem. In particular, we model review dataset as a HIN in which reviews are connected through different node types (such as features and users). To evaluate the proposed solution, we used two sample review datasets from Yelp and Amazon websites.

1.2 Motivation

This paper proposed to use duplicate detection and classification to detect review spam. Our preliminary experiments showed promising results. Our future work will focus on improving the accuracy and detecting more sophisticated spam reviews. Based on our observations, defining two views for features (review-user and behavioral-linguistic), the classified features as review behavioral have more weights and yield better performance on spotting spam reviews in both semi-supervised and unsupervised approaches. In addition, we demonstrate that using different supervisions such as 1%, 2.5% and 5% or using an unsupervised approach, make no noticeable variation on the performance of our approach. We observed that feature weights can be added or removed for labeling and hence time complexity can be scaled for a specific level of accuracy. As the result of this weighting step, we can use fewer features with more weights to obtain better accuracy with less time complexity. In addition, categorizing features in four major categories (review-behavioral, user-behavioral, review linguistic, user-linguistic), helps us to understand how much each category of features is contributed to spam detection.

1.3 Problem Definition

Social Media websites play a main role in information propagation which is considered as an important source for producers in their advertising operations as well as for customers in selecting products and services. People mostly believe on the written reviews in their decision-making processes, and positive/negative reviews encouraging/discouraging them in their selection of products and services. These reviews thus have become an important factor in success of a business while positive reviews can bring benefits for a company, negative reviews can potentially impact credibility and cause economic losses. The fact that anyone with any identity can leave comments as reviews provides a tempting opportunity for spammers to write fake reviews designed to mislead users' opinion. These misleading reviews are then multiplied by the sharing function of social media and propagation over the web. The reviews written to change users' perception of how good a product or a service are considered as spam, and are often written in exchange for money.

Disadvantages:

1. There is no information filtering concept in social network.
2. People believe on the written reviews in their decision-making processes, and positive/negative reviews encouraging/discouraging them in their selection of products and services.
3. Anyone can access application through registration and gives feedbacks as reviews for spammers to misguide other user's opinion.
4. Less accuracy.
5. More time complexity.

II. LITERATURE SURVEY

The paper [1] represents the pairwise features are first explicitly utilized to detect group colluders in online product review spam campaigns, which can reveal collusions in spam campaigns from a more fine-grained perspective. A novel detecting framework named FraudInformer is proposed to cooperate with the pairwise features which are intuitive and unsupervised. Advantages are: Pairwise features can be more robust model for correlating colluders. Manipulate perceived reputations of the targets for their best interests. To rank all the reviewers in the website globally so that top-ranked ones are more likely to be colluders. Disadvantages are: Difficult problem to automate.

The paper [2] builds a network of reviewers appearing in different bursts and model reviewers and their co-occurrence in bursts as a Markov Random Field (MRF), and employ the Loopy Belief Propagation (LBP) method to infer whether a reviewer is a spammer or not in the graph. A novel evaluation method to evaluate the detected spammers automatically using supervised classification of their reviews. Advantages are: High accuracy. The proposed method is effective. To detect review spammers in review bursts. Detect spammers automatically. Disadvantages are: a generic framework is not used for detect spammers.

In paper [3], the challenges are: The detection of fraudulent behaviors, assessing the trustworthiness of review sites, since some may have policies that enable misbehavior, and creating effective review aggregation solutions. The TrueView score, in three different variants, as a proof of concept that the synthesis of multi-site

reviews can provide important and usable information to the end user. Advantages are: Develop novel features capable of identifying cross-site discrepancies effectively. A hotel identity-matching method has 93% accuracy. Enable the site owner to detect misbehaving hotels. Enable the end user to trusted reviews. Disadvantages are: Difficult problem to automate.

The paper [4] describes unsupervised anomaly detection techniques over user behavior to distinguish potentially bad behavior from normal behavior. To detect diverse attacker strategies fake, compromised, and colluding Facebook identities with no a priori labeling while maintaining low false positive rates. Advantages are: Anomaly detection technique to effectively identify anomalous likes on Facebook ads. Achieves a detection rate of over 66% (covering more than 94% of misbehavior) with less than 0.3% false positives. Disadvantages are: The attacker is trying to drain the budget of some advertiser by clicking on ads of that advertiser.

In [5] paper, a collective classification algorithm called Multi-typed Heterogeneous Collective Classification (MHCC) and then extends it to Collective Positive and Unlabeled learning (CPU).The proposed models can markedly improve the F1 scores of strong baselines in both PU and non-PU learning settings. Advantages are: Proposed models can markedly improve the F1 scores of strong baselines in both PU and non-PU learning settings. Models only use language independent features; they can be easily generalized to other languages. Detects a large number of potential fake reviews hidden in the unlabeled set. Disadvantages are: Fake reviews hiding in the unlabel reviews that Dianping's algorithm did not capture. The ad-hoc labels of users and IPs used in MHCC may not be very accurate as they are computed from labels of neighboring reviews.

The paper [6] elaborates two distinct methods of reducing feature subset size in the review spam domain. The methods include filter-based feature rankers and word frequency based feature selection. Advantages are: The first method is to simply select the words which appear most often in the text. Second method can use filter based feature rankers (i.e. Chi-Squared) to rank features and then select the top ranked features. Disadvantages are: There is not a one size fits all approach that is always better.

In [7] paper, providing an efficient and effective method to identify review spammers by incorporating social relations based on two assumptions that people are more likely to consider reviews from those connected with them as trustworthy, and review spammers are less likely to maintain a large relationship network with normal users. Advantages are: The proposed trust-based prediction achieves a higher accuracy than standard CF method. To overcome the sparsity problem and compute the overall trustworthiness score for every user in the system, which is used as the spamicity indicator. Disadvantages are: Review dataset required.

Sr. No.	**Author, Title and Journal Name**	**Advantages**	**Disadvantage**	**Refer Points**
1	Ch. Xu and J. Zhang. Combating product review spam campaigns via multiple heterogeneous pairwise features. In SIAM International Conference on Data Mining, 2014.	1. Pairwise features can be more robust model for correlating colluders. 2. To manipulate perceived reputations of the targets for their best interests. 3. To rank all the reviewers in the website globally so that top-ranked ones are more likely to be colluders.	1. Difficult problem to automate.	The pairwise features are first explicitly utilized to detect group colluders in online product review spam campaigns, which can reveal collusions in spam campaigns from a more fine-grained perspective. A novel detecting framework named FraudInformer is proposed to cooperate with the pairwise features which are intuitive and unsupervised.
2	G. Fei, A. Mukherjee, B. Liu, M. Hsu, M. Castellanos, and R. Ghosh. Exploiting burstiness in reviews for review spammer detection. In ICWSM, 2013.	1. High accuracy. 2. The proposed method is effective. 3. To detect review spammers in review bursts. 4. To detect spammers automatically.	1. a generic framework is not used for detect spammers	To build a network of reviewers appearing in different bursts and model reviewers and their co-occurrence in bursts as a Markov Random Field (MRF), and employ the Loopy Belief Propagation (LBP) method to infer whether a reviewer is a spammer or not in the graph. A novel evaluation method to evaluate the detected spammers automatically using supervised classification of their reviews.
3	A. j. Minnich, N. Chavoshi, A. Mueen, S. Luan, and M. Faloutsos. Trueview: Harnessing the power of multiple review sites. In ACM WWW, 2015.	1. Develop novel features capable of identifying cross-site discrepancies effectively. 2. A hotel identity-matching method with 93% accuracy.	1. Difficult problem to automate.	In this paper, the challenges are: The detection of fraudulent behaviors, assessing the trustworthiness of review sites, since some may have policies that enable misbehavior, and creating effective review aggregation

		3. Enable the site owner to detect misbehaving hotels. 4. Enable the end user to trusted reviews.		solutions. The TrueView score, in three different variants, as a proof of concept that the synthesis of multi-site reviews, can provide important and usable information to the end user.
4	B. Viswanath, M. Ahmad Bashir, M. Crovella, S. Guah, K. P. Gummadi, B. Krishnamurthy, and A. Mislove. Towards detecting anomalous user behavior in online social networks. In USENIX, 2014.	1. Anomaly detection technique to effectively identify anomalous likes on Facebook ads. 2. Achieves a detection rate of over 66% (covering more than 94% of misbehavior) with less than 0.3% false positives.	1. The attacker is trying to drain the budget of some advertiser by clicking on ads of that advertiser.	Unsupervised anomaly detection techniques over user behavior to distinguish potentially bad behavior from normal behavior. To detect diverse attacker strategies fake, compromised, and colluding Facebook identities with no a priori labeling while maintaining low false positive rates.
5	H. Li, Z. Chen, B. Liu, X. Wei, and J. Shao. Spotting fake reviews via collective PU learning. In ICDM, 2014.	1. Proposed models can markedly improve the F1 scores of strong baselines in both PU and non-PU learning settings. 2. Models only use language independent features; they can be easily generalized to other languages. 3. Detects a large number of potential fake reviews hidden in the unlabeled set.	1. Fake reviews hiding in the unlabel reviews that Dianping's algorithm did not capture. 2. The ad-hoc labels of users and IPs used in MHCC may not be very accurate as they are computed from labels of neighboring reviews.	In this paper, a collective classification algorithm called Multi-typed Heterogeneous Collective Classification (MHCC) and then extends it to Collective Positive and Unlabeled learning (CPU). The proposed models can markedly improve the F1 scores of strong baselines in both PU and non-PU learning settings.
6	M. Crawford, T. M. Khoshgoftaar, and J.D. Prusa. ReducingFeature set Explosionto FaciliateReal-World Review SpamDetection. InProceedingof 29th InternationalFlorida ArtificialIntelligence Research SocietyConference.2016.	1. The first method is to simply select the words which appear most often in the text. 2. Second method can use filter based feature rankers (i.e. Chi-Squared) to rank features and then select the top ranked features.	1. There is not a one size fits all approach that is always better.	In this paper, consider two distinct methods of reducing feature subset size in the review spam domain. The methods include filter-based feature rankers and word frequency based feature selection.
7.	H. Xue, F. Li, H. Seo, and R. Pluretti. Trust- Aware Review Spam Detection. IEEE Trustcom/ISPA. 2015.	1. The proposed trust-based prediction achieves a higher accuracy than standard CF method. 2. To overcome the sparsity problem and compute the overall trustworthiness score for every user in the system, which is used as the spamicity indicator.	1. Review dataset required.	In this paper, providing an efficient and effective method to identify review spammers by incorporating social relations based on two assumptions that people are more likely to consider reviews from those connected with them as trustworthy, and review spammers are less likely to maintain a large relationship network with normal users.

III. EXISTING SYSTEM APPROACH

Online Social Media websites play a main role in information propagation which is considered as an important source for producers in their advertising operations as well as for customers in selecting products and services. People mostly believe on the written reviews in their decision-making processes, and positive/negative reviews encouraging/discouraging them in their selection of products and services. These reviews thus have become an important factor in success of a business while positive reviews can bring benefits for a company, negative reviews can potentially impact credibility and cause economic losses. The fact that anyone with any identity can leave comments as reviews provides a tempting opportunity for spammers to write fake reviews designed to mislead users' opinion. These misleading reviews are then multiplied by the sharing function of social media and propagation over the web. The reviews written to change users' perception of how good a product or a service are considered as spam, and are often written in exchange for money.

Disadvantages:
1. There is no information filtering concept in online social network.
2. People believe on the written reviews in their decision-making processes, and positive/negative reviews encouraging/discouraging them in their selection of products and services.
3. Anyone create registration and gives comments as reviews for spammers to write fake reviews designed to misguide users' opinion.
4. Less accuracy.
5. More time complexity.

IV. PROPOSED SYSTEM APPROACH

The proposed framework is to model a given review dataset as a Heterogeneous Information Network (HIN) and to map the problem of spam detection into a HIN classification problem. In particular, we model review dataset as a HIN in which reviews are connected through different node types (such as features and users). A weighting algorithm is then employed to calculate each feature's importance (or weight). These weights are utilized to calculate the final labels for reviews using both unsupervised and supervised approaches. Based on our observations, defining two views for features (review-user and behavioral-linguistic), the classified features as review behavioral have more weights and yield better performance on spotting spam reviews in both semi-supervised and unsupervised approaches. The feature weights can be added or removed for labeling and hence time complexity can be scaled for a specific level of accuracy. Categorizing features in four major categories (review-behavioral, user-behavioral, review-linguistic, user-linguistic), helps us to understand how much each category of features is contributed to spam detection.
1. NetSpam framework that is a novel network based approach which models review networks as heterogeneous information networks.
2. A new weighting method for spam features is proposed to determine the relative importance of each feature and shows how effective each of features are in identifying spams from normal reviews.
3. NetSpam improves the accuracy compared to the state-of-the art in terms of time complexity, which highly depends to the number of features used to identify a spam review.

The general concept of our proposed framework is to model a given review dataset as a Heterogeneous Information Network and to map the problem of spam detection into a HIN classification problem. In particular, we model review 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. Based on our observations defining two views for features.

Advantages:
1. To identify spam and spammers as well as different type of analysis on this topic.
2. Written reviews also help service providers to enhance the quality of their products and services.
3. To identify the spam user using positive and negative reviews in online social media.
4. To display only trusted reviews to the users.

V. SYSTEM ARCHITECTURE

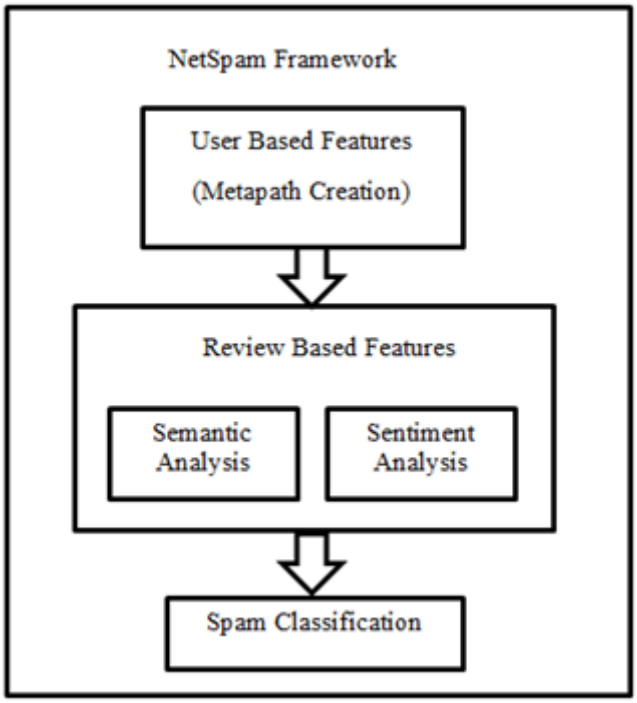

Fig No 01 System Architecture

5.1 Mathematical model for the low-level design (module-wise)

Spam Features:

User-Behavioral (UB) based features:

Burstiness: Spammers, usually write their spam reviews in short period of time for two reasons: first, because they want to impact readers and other users, and second because they are temporal users, they have to write as much as reviews they can in short time.

$$x_{BST}(i) = \begin{cases} 0 & (L_i - F_i) \notin (0, \tau) \\ 1 - \frac{L_t - F_t}{\tau} & (L_i - F_i) \in (0, \tau) \end{cases} \quad (1)$$

Where,

$L_i - F_i$ describes days between last and first review for τ = 28.

Users with calculated value greater than 0.5 take value 1 and others take 0.

User-Linguistic (UL) based features:

Average Content Similarity, Maximum Content Similarity: Spammers, often write their reviews with same template and they prefer not to waste their time to write an original review. In result, they have similar reviews. Users have close calculated values take same values (in [0; 1]).

Review-Behavioral (RB) based features:

- Early Time Frame: Spammers try to write their reviews a.s.a.p., in order to keep their review in the top reviews which other users visit them sooner.

$$x_{ETF}(i) = \begin{cases} 0 & (L_i - F_i) \notin (0, \delta) \\ 1 - \frac{L_t - F_t}{\delta} & (L_i - F_i) \in (0, \delta) \end{cases} \quad (2)$$

  Where,

  $L_i - F_i$ denotes days specified written review and first written review for a specific business. We have also δ = 7. Users with calculated value greater than 0.5 takes value 1 and others take 0.

- Rate Deviation using threshold: Spammers, also tend to promote businesses they have contract with, so they rate these businesses with high scores. In result, there is high diversity in their given scores to different businesses which is the reason they have high variance and deviation.

$$x_{DEV}(i) = \begin{cases} 0 & Otherwise \\ 1 - \frac{r_{ij} - avg_{e \in E*j} r(e)}{4} > \beta_1 \end{cases} \quad (3)$$

Where,
β_1 is some threshold determined by recursive minimal entropy partitioning. Reviews are close to each other based on their calculated value, take same values (in [0; 1)).

Review-Linguistic (RL) based features:
Number of first Person Pronouns, Ratio of Exclamation Sentences containing '!': First, studies show that spammers use second personal pronouns much more than first personal pronouns. In addition, spammers put '!' in their sentences as much as they can to increase impression on users and highlight their reviews among other ones. Reviews are close to each other based on their calculated value, take same values (in [0; 1]).

VI. RESULT AND ANALYSIS

Experimental evaluation results shows the Tripadvisor hotel review dataset with higher percentage of spam reviews have better performance because when fraction of spam reviews increases, probability for a review to be a spam review increases and as a result more spam reviews will be labeled as spam reviews. The results of the dataset show all the four behavioral features are ranked as first features in the final overall weights. The Fig.2 graph shows the NetSpam framework features for the dataset have more weights and features for Review-based dataset stand in the second position. Third position belongs to User-based dataset and finally Item-based dataset has the minimum weights (for at least the four features with most weights).

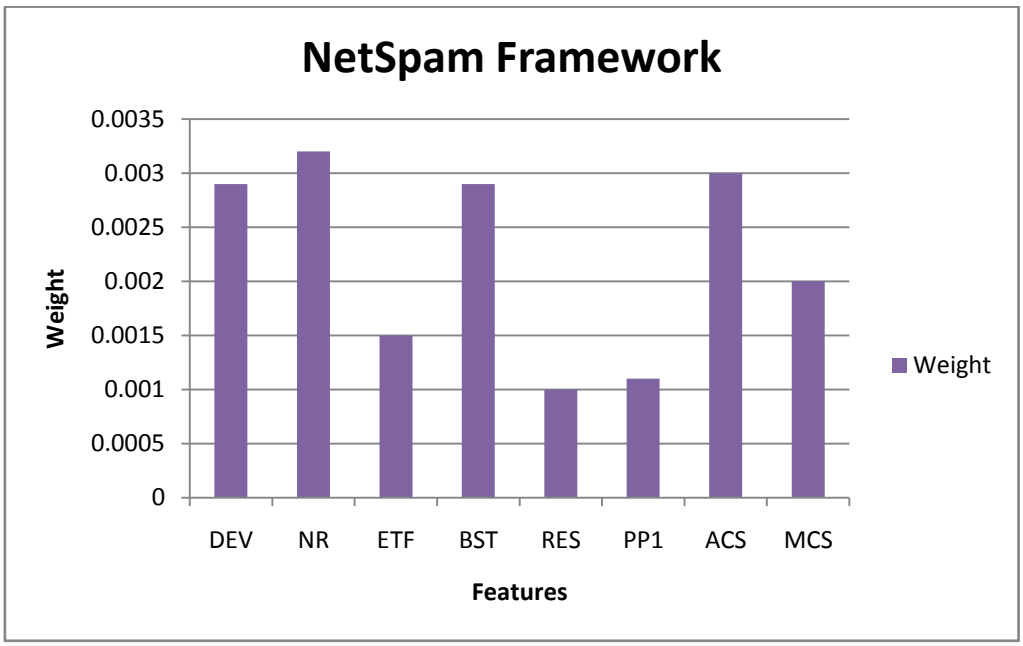

Fig.2 Feature weights for NetSpam Framework

TABLE I Weights of all features

Features	Weight
DEV	0.0029
NR	0.0032
ETF	0.0015
BST	0.0029
RES	0.001
PP1	0.0011
ACS	0.003
MCS	0.002

CONCLUSION

This investigation presents a novel spam detection system in particular NetSpam in view of a metapath idea and another graph based strategy to name reviews depending on a rank-based naming methodology. The execution of the proposed structure is assessed by utilizing review datasets. Our perceptions demonstrate that ascertained weights by utilizing this metapath idea can be exceptionally powerful in recognizing spam surveys and prompts a superior execution. Furthermore, we found that even without a prepare set, NetSpam can figure the significance of each element and it yields better execution in the highlights' expansion procedure, and

performs superior to anything past works, with just few highlights. In addition, in the wake of characterizing four fundamental classifications for highlights our perceptions demonstrate that the review behavioral classification performs superior to anything different classifications, regarding AP, AUC and in the ascertained weights. The outcomes likewise affirm that utilizing diverse supervisions, like the semi-administered strategy, have no detectable impact on deciding the vast majority of the weighted highlights, similarly as in various datasets. Contribution part in this project, for user when searches query he will get the top-k hotel lists as well as one recommendation hotel by using personalized recommendation algorithm.

REFERENCES

[1]. J. Donfro, A whopping 20 % of yelp reviews are fake. http://www.businessinsider.com/20-percent-of-yelp-reviews-fake-2013-9. Accessed: 2015-07-30.
[2]. M. Ott, C. Cardie, and J. T. Hancock. Estimating the prevalence of deception in online review communities. In ACM WWW, 2012.
[3]. M. Ott, Y. Choi, C. Cardie, and J. T. Hancock. Finding deceptive opinion spam by any stretch of the imagination.In ACL, 2011.
[4]. Ch. Xu and J. Zhang. Combating product review spam campaigns via multiple heterogeneous pairwise features. In SIAM International Conference on Data Mining, 2014.
[5]. N. Jindal and B. Liu. Opinion spam and analysis. In WSDM, 2008.
[6]. F. Li, M. Huang, Y. Yang, and X. Zhu. Learning to identify review spam. Proceedings of the 22nd International Joint Conference on Artificial Intelligence; IJCAI, 2011.
[7]. G. Fei, A. Mukherjee, B. Liu, M. Hsu, M. Castellanos, and R. Ghosh. Exploiting burstiness in reviews for review spammer detection. In ICWSM, 2013.
[8]. A. j. Minnich, N. Chavoshi, A. Mueen, S. Luan, and M. Faloutsos. Trueview: Harnessing the power of multiple review sites. In ACM WWW, 2015.
[9]. B. Viswanath, M. Ahmad Bashir, M. Crovella, S. Guah, K. P. Gummadi, B. Krishnamurthy, and A. Mislove. Towards detecting anomalous user behavior in online social networks. In USENIX, 2014.
[10]. H. Li, Z. Chen, B. Liu, X. Wei, and J. Shao. Spotting fake reviews via collective PU learning. In ICDM, 2014.
[11]. L. Akoglu, R. Chandy, and C. Faloutsos. Opinion fraud detection in online reviews bynetwork effects. In ICWSM, 2013.
[12]. R. Shebuti and L. Akoglu. Collective opinion spam detection: bridging review networksand metadata. In ACM KDD, 2015.
[13]. S. Feng, R. Banerjee and Y. Choi. Syntactic stylometry for deception detection. Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics: Short Papers; ACL, 2012.
[14]. N. Jindal, B. Liu, and E.-P. Lim. Finding unusual review patterns using unexpected rules. In ACM CIKM, 2012.