Multiple criteria fake reviews detection based on spammers' indicators within the belief function theory

Malika Ben Khalifa, Zied Elouedi, Eric Lefevre

Multiple criteria fake reviews detection based on spammers' indicators within the belief function theory

Malika Ben Khalifa[1,2], Zied Elouedi[1], and Eric Lefèvre[2]

[1] Université de Tunis, Institut Supérieur de Gestion de Tunis, LARODEC, Tunisia
malikabenkhalifa2@gmail.com, zied.elouedi@gmx.fr
[2] Univ. Artois, EA 3926, Laboratoire de Génie Informatique et d"Automatique de l"Artois (LGI2A), Béthune, F-62400, France
eric.lefevre@univ-artois.fr

Abstract. E-reputation becomes one of the most important keys of success for companies and brands. It is mainly based on the online reviews which significantly influence consumer purchase decisions. Therefore, in order to mislead and artificially manipulate costumers' perceptions about products or services, some dealers rely on spammers who post fake reviews to exaggerate the advantages of their products and defame rival's reputation. Hence, fake reviews detection becomes an essential task to protect online reviews, maintain readers' confidence and to ensure companies fair competition. In this way, we propose a new method based on both the reviews given to multiple evaluation criteria and the reviewers' behaviors to spot spam reviews. This approach deals with uncertainty in the different inputs thanks to the belief function theory. Our method shows its performance in fake reviews detection while testing with two large real-world review data-sets from Yelp.com.

Keywords: Online reviews, Fake reviews, Spammers, Multi-criteria evaluation, Uncertainty, Belief function theory.

1 Introduction

Nowadays, online reviews are one of the most valuable sources of information for customers. Moreover, they are considerate as the pillars on which companies' reputation is built. Most of consumers believe in checking the reviews given to a product or service before deciding to purchase it. Therefore, companies with high number of positive reviews are lucked to attract a huge number of new consumers and consequently they will achieve significant financial gains. However, negative reviews or lowest rating reviews cause financial losses. Driven by the profit, some companies and brands pay spammers to post fake positive reviews on their own product in order to enhance their e-reputation, not only that but also they try to damage competitors' reputation by posting negative reviews on their products. Consequently, detection of opinion spam actually becomes more and more big concern to protect online opinions, to gain consumer trust and

to maintain companies' fair competition. For this reason, in the last years, several methods have been proposed trying to distinguish between the trustful and deceptive reviews. All the first studies rely on the review content using the linguistic aspects and feeling as well as readability and subjectivity [5]. Moreover, other techniques based on the individual words extracted from the review text as features [9], while some others are based on the syntactic and lexical features. It is important to mention that most of the methods based only on the review content can not successfully detect fake reviews cause of the lack of any distinguishing words that can give a definitive clue for classification of reviews as real or fake. Accordingly, detecting spammers can improve spotting spam review, since spammers generally share the same profile history and activity patterns. Hence, we notice the existing of various spammer detection methods in which the graph-theory have been used and most of them have shown promising results [18]. Moreover, other methods [7, 11, 13] are based on the different features extracted from the reviewer characteristics and behavioral. Furthermore, relying on both spam review detection and spammer detection when analyzing their behaviors is more effective solution for detecting review spam than either approach alone. In this way, we mention works in [8, 15], that exploit both relational data and metadata of reviewers and reviews. Results prove that this kind of methods outperform all others.
Although the fake reviews detection is an uncertain problem, no one of these previous works is able to manage uncertainty in the reviews. However, we have proposed some preliminary related works dealing with the uncertainty [1, 2] but these approaches rely only on the review information. In this paper, we propose a novel approach that distinguishes between fake and genuine reviews while dealing with uncertainty in both the review and the reviewer information. As some reviewers prefer judge services or products through different evaluation criteria, our method deals with the different review rating criteria and analyzes the reviewers behaviors under the belief function framework, chosen thanks to its flexibility in representing and managing different types of imperfection. The rest of this paper is organized as follows: In Section 2, we remind the basic concepts of the belief function theory. Then, we elucidate our proposed approach in Section 3. Section 4 discusses the experimental study. Finally, we conclude in Section 5.

2 Belief Function Theory

The belief function theory is one of the useful theories that handles uncertain knowledge. It was introduced by Shafer [16] as a model to manage beliefs. The frame of discernment Ω is a finite and exhaustive set of different events associated to a given problem. 2^{Ω} is the power set of Ω that contains all possible hypotheses. A basic belief assignment (*bba*) or a belief mass is defined as a function from 2^{Ω} to $[0, 1]$ that represents the degree of belief given to an element A such that: $\sum_{A \subseteq \Omega} m^{\Omega}(A) = 1$.
A focal element A is a set of hypotheses with positive mass value $m^{\Omega}(A) > 0$. Moreover, we underline some special cases of *bba*'s:

- The certain *bba* represents the state of total certainty and it is defined as follows: $m^{\Omega}(\{\omega_i\}) = 1$ and $\omega_i \in \Omega$.
- Simple support function: In this case, the *bba* focal elements are $\{A, \Omega\}$. A simple support function is defined as the following equation:

$$m^{\Omega}(X) = \begin{cases} w & \text{if } X = \Omega \\ 1 - w & \text{if } X = A \text{ for some } A \subset \Omega \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

  where A is the focus and $w \in$ [0,1].

Moreover, the discounting operation [12] allows us to update experts beliefs by taking into consideration their reliability through the degree of trust $(1 - \alpha)$ given to each expert with $\alpha \in [0, 1]$ is the discount rate.
Accordingly, the discounted *bba*, noted ${}^{\alpha}m^{\Omega}$, m^{Ω} becomes:

$$\begin{cases} {}^{\alpha}m^{\Omega}(A) = (1 - \alpha)m^{\Omega}(A) & \forall A \subset \Omega, \\ {}^{\alpha}m^{\Omega}(\Omega) = \alpha + (1 - \alpha)m^{\Omega}(\Omega). \end{cases} \tag{2}$$

Several combination rules have been proposed in the framework of belief function to aggregate a set of *bba*'s provided by pieces of evidence from different experts. Let m_1^{Ω} and m_2^{Ω} two *bba*'s modeling two distinct sources of information defined on the same frame of discernment Ω. In what follows, we elucidate the combination rules related to our approach.

1. *Conjunctive rule*: It was settled in [17], denoted by $\textcircled{\cap}$ and defined as:

$$m_{\textcircled{\cap}}^{\Omega}(A) = m_1^{\Omega} \textcircled{\cap} m_2^{\Omega}(A) = \sum_{B \cap C = A} m_1^{\Omega}(B) m_2^{\Omega}(C), \qquad \forall B, C \subseteq \Omega$$

2. *Dempster's rule of combination*: This combination rule is a normalized version of the conjunctive rule [4]. It is denoted by $\oplus$ and defined as:

$$m_{\oplus}^{\Omega}(A) = m_1^{\Omega} \oplus m_2^{\Omega}(A) = \begin{cases} \frac{m_1^{\Omega} \textcircled{\cap} m_2^{\Omega}(A)}{1 - m_1^{\Omega} \textcircled{\cap} m_2^{\Omega}(\emptyset)} & \text{if} \quad A \neq \emptyset, \forall A \subseteq \Omega, \\ 0 & otherwise. \end{cases} \tag{3}$$

3. *The combination with adapted conflict rule* $(CWAC)$**:** This combination [6] is an adaptive weighting between the two previous combination rules acting like the conjunctive rule if *bbas* are opposite and as the Dempster rule otherwise. They use the notion of dissimilarity that is obtained through a distance measure, to ensure this adaptation between all sources.

The CWAC is formulated as follows:

$$m_{\textcircled{\leftrightarrow}}^{\Omega}(A) = (\textcircled{\leftrightarrow} m_i^{\Omega})(A) = D_{max} m_{\textcircled{\cap}}^{\Omega}(A) + (1 - D_{max}) m_{\oplus}^{\Omega}(A) \tag{4}$$

where D_{max} is the maximal value of all the distances, it can be used to find out if at least one of the sources is opposite to the others and thus it may be defined by: $D_{max} = max[d(m_i^{\Omega}, m_j^{\Omega})]$ where $i \in [1, M]$, $j \in [1, M]$, M is the total number of

mass functions and $d(m_i^\Omega, m_j^\Omega)$ is the distance measure proposed by Jousselme [10]: $d(m_1^\Omega, m_2^\Omega) = \sqrt{\frac{1}{2}(m_1^\Omega - m_2^\Omega)^t D(m_1^\Omega - m_2^\Omega)}$, where D is the Jaccard index defined by: $D(E,F) = \begin{cases} 1 & \text{if} \quad E = F = \emptyset, \\ \frac{|E \cap F|}{|E \cup F|} & \forall E, F \in 2^\Omega \backslash \emptyset \end{cases}$

Frequently, we need to fuse two $bba's$ $m_1^{\Omega_1}$ and $m_2^{\Omega_2}$ that have not the same frame of discernment. So, we apply the vacuous extension of the belief function which extend the frames of discernment Ω_1 and Ω_2, corresponding to the mass functions $m_1^{\Omega_1}$ and $m_2^{\Omega_2}$, to the product space $\Omega = \Omega_1 \times \Omega_2$. The vacuous extension operation, denoted by $\uparrow$ and defined such that:
$m^{\Omega_1 \uparrow \Omega_1 \times \Omega_2}(B) = m^{\Omega_1}(A) \quad if \quad B = A \times \Omega_2$ where $A \subseteq \Omega_1$, $B \subseteq \Omega_1 \times \Omega_2$.
It transforms each mass to the cylindrical extension B to $\Omega_1 \times \Omega_2$.
To determine relation between two disjoint frames of discernment Ω_1 and Ω_2, the multi-valued mapping may be used. This operation denoted τ, allows us to join together two different frame of discernment the subsets $B \subseteq \Omega_2$ that can match through τ to be a subset $A \subseteq \Omega_1$: $m_\tau^{\Omega_1}(A) = \sum_{\tau(B)=A} m^{\Omega_2}(B)$
The belief function framework offers various solutions to ensure the decision making. We present the pignistic probabilities, used in our work, denoted by $BetP$ and defined as: $BetP(B) = \sum_{A \subseteq \Omega} \frac{|A \cap B|}{|A|} \frac{m^\Omega(A)}{(1 - m^\Omega(\emptyset))} \quad \forall\ B \in \Omega$.

3 Multiple criteria fake reviews detection based on spammers' indicators within the belief function theory

In this following section, we elucidate our novel approach which aims to detect fake reviews. Our method relies on both the reviews and the reviewers information, since gathering behavioral evidence of spammers in more efficient than just identifying review spam only. Moreover, our proposed approach is divided on three parts: Firstly, we deal with the ratings given to various evaluation criteria, to judge a service or a product, in order to determine the reviewers' opinion trustfulness through their degree with compatibility with all others' opinions, this part is based on our previous work in [2]. Secondly, trying to obtain more preferment detection, we propose to rely on an other previous work in which we model the reviewers' trustworthiness by analyzing the reviewers' behaviors [3]. For that, we adopt the belief function theory to handle uncertainty in the various imprecise reviews and the imperfect reviewers' information. To enhance the detection performance, we combine both the reviewer opinion and the reviewer trustworthiness modeling by mass functions in the third part. These three parts are detailed in the following subsections in which we consider a dataset of N reviewers and q evaluation criteria. Each reviewer R_i evaluates a product or a service by giving a rating vote between 1 and 5 stars to each criterion C_j.

3.1 Modeling the reviewer's opinion trustworthiness

Considering the rating reviews given to different evaluation criteria as inputs, where each vote V_{ij} provided to each criterion C_j where j in $\{1...q\}$ and q the

number of evaluation criteria. Therefore, we propose to model uncertainty in these rating reviews by representing each vote into mass function $m_{ik}^{\Omega_j}$ with Ω_j $=\{1,2,3,4,5\}$ where each element represents the rating number given by each reviewer R_i.

3.1.1 Modeling the uncertain opinion

We think that the reviewer gives always imprecise vote to one value close. Hence, we model this uncertainty by considering, the vote, the vote+1, the vote-1 denoted k, for each rating review (i.e vote) given to each criterion. Hence, we transform the uncertain vote into a mass functions (i.e $bba's$) such that: $m_{ik}^{\Omega_j}(\{k\}) = 1$ where $k \in \{V_i, V_{i+1}, V_{i-1}\}$.

Then, we propose to take into consideration the reliability degree of each vote V_{ij} based on its similarity with all others' vote provided to the same criterion C_j. We also take into account the difference between the vote given by the reviewer and these modeled values ($\{V_i, V_{i+1}, V_{i-1}\}$) denoted by k. For this we apply double discounting in which we transform the mass functions on simple support system. After that, we aggregate the discounted *bba* representing the given vote using the Dempster rule (Eq.3) in order to model each uncertain vote given to each criterion by one global mass function $m_i^{\Omega_j}$ with $i = 1, \ldots, N$ and $j = 1, \ldots, q$.

Moreover, we propose to model the whole reviewer's opinion in one joint *bba*, for that we have to apply the following steps:

- Creating Ω_c as the global frame of discernment relative to all criteria which represents the cross product of the different frames Ω_j denoted by: $\Omega_c = \Omega_1 \times \Omega_2 \times \ldots \times \Omega_q$.
- Extending the different $bba's$ $m_i^{\Omega_j}$ to the global frame Ω_c to get new $bba's$ $m_i^{\Omega_j \uparrow \Omega_c}$.
- Combining the extended $bba's$ using the Dempster rule of combination. $m_i^{\Omega_c} = m_i^{\Omega_1 \uparrow \Omega_c} \oplus m_i^{\Omega_2 \uparrow \Omega_c} \oplus \ldots \oplus m_i^{\Omega_q \uparrow \Omega_c}$. This *bba* represents the reviewer's opinion given by the different rating review criteria.

3.1.2 Measure the compatibility between the reviewer opinion and all the others' one

In order to evaluate the opinion provided by each reviewer R_i through various criteria C_j, we compare it with all others reviewers' opinions. For that, firstly we aggregate all the others review rating given to the same criterion using the CWAC combination rule to obtain one *bba*, modeling the whole rating reviews given each criterion except the current one. As a consequence, we obtain q (number of evaluation criterion) $bba's$ $m_{ic}^{\Omega_j}$. Thus, we combine them to model all reviewers' opinions except the current one in one joint *bba*. To achieve this, we firstly extend them to the global frame of criteria Ω_c to get $m_{ic}^{\Omega_j \uparrow \Omega_c}$. After that, we aggregate the extend $bba's$ through the Dempster rule of combination. Subsequently, for each reviewer we measure the distance between his provided opinion modeled by $m_i^{\Omega_c}$ and all the others reviewers' opinions represented by $m_{ic}^{\Omega_c}$ using the distance of Jousselme.

3.1.3 Modeling the reviewer opinion into trustful or not trustful

Since the calculated distance elucidates the average opinion rating deviation from the other reviewers' opinion which one of the most important spam indicator. That's why, more the distance decreases more the given opinion is considerate as trustful. Thus, we propose to transform each distance into new *bba* under $\Theta = \{t, \bar{t}\}$ (t for trustful and $\bar{t}$ for not trustful).
In this part, we successfully model the whole reviewers' opinion trustworthiness by mass function m_i^{Θ} under the frame of discernment $\Theta = \{t, \bar{t}\}$.

3.2 Modeling the reviewer spamicity

The average rating deviation from the others rating is considerate as an important indicators in spam review detection field. Despite that, spammers try to mislead readers and the usually post a lot of despite reviews to dominate the majority of the given opinions. Accordingly, it is essential to have recourse to the reviewers' spamicity in order to reinforce the spam reviews detection. In this way, we propose to model uncertainty in the different reviewers information while relying on the spammers indicators. We represent each reviewer R_i by two mass functions namely; the reviewer reputation $m_{RR_i}^{\Omega_S}$ and the second one is to model the reviewer helpfulness $m_{RH_i}^{\Omega_S}$ with $\Omega_S = \{S, \overline{S}\}$ where S is spammer and $\overline{S}$ is not spammer.

3.2.1 Modeling the reviewer reputation

Usually, the innocent reviewers post their opinion when they have already bought new products or used new services. Therefore, their reviews are generally dispersed over time interval and depend on the number of used products or services. However, the spammers post enormous reviews to some particular products in short interval, two or three days (more used in the spammer review detection field), to overturn the majority of the given reviews. Consequently, we construct the reviewer reputation through these two spammers' indicators. Hence, we propose to check the reviewing history of each reviewer $Hist_{R_i}$ contained all past reviews given by each reviewer R_i to n discrete products or services. Each reviewer average proliferation is calculated as follows: $AvgP(R_i) = \frac{Hist_{R_i}}{n}$
We assume that if $AvgP(R_i) > 3$, the reviewer is considered as a potential spammer since generally ordinary reviewers do not give more than three reviews per product [13]. Accordingly, the reviewer reputation is then represented by a certain *bba* as follows:

$$m_{RR_i}^{\Omega_S}(\{S\}) = 1 \text{ else } m_{RR_i}^{\Omega_S}(\{\overline{S}\}) = 1$$

Moreover, we check if the reviews are given in a short time of interval or are distributed all along the reviewing history.
For that, we measure the brust spamicity degree denoted δ_i to weaken each reviewer reputation *bba* by its corresponding reliability degree using the discounting operation. Consequently, we find the discounted *bba* ${}^{\delta_i}m_{RR_i}^{\Omega_S}$ which presents the reviewer reputation based on both the reviewer's average proliferation and the brust spamicity.

3.2.2 Modeling the reviewer helpfulness

The reviewer helpfulness is considerate as important spammer indicators. Thus, we extract the Number of Helpful Reviews (NHR) associated to each reviewer in order to check if the reviewer post helpful reviews or unhelpful one to mislead readers. Hence, if the reviewer is suspicious to be spammer ($NHR_i = 0$), thus we model the reviewer helpfulness by a certain *bba* as follows:

$$m^{\Omega_S}_{RH_i}(\{S\}) = 1 \text{ else } m^{\Omega_S}_{RH_i}(\{\overline{S}\}) = 1$$

We weaken the reviewer helpfulness mass by the non helpfulness degree for each reviewer R_i denoted by λ_i in order to not treat all the reviewers who give helpful reviews in the same way. Thus, we apply the discounting operation in order to transform the *bba* into a simple support function ${}^{\lambda_i}m^{\Omega_S}_{RH_i}$ to take into consideration the helpfulness degree. Moreover, spammers usually post extreme ratings [13], either highest (5 stars) or lowest (1 star), in order to achieve their objective and dominate the average rating score of products or services. Nonetheless, the innocent reviewers are always not fully satisfied or dissatisfied by the tried products and services. Thus, they not usually post extreme rating. Hence, despite the fact that the reviewer has helpful reviews if they are crowded with extreme rating, his probabilities of being genuine reviewer will assuredly reduce.
In order to take this indicator into consideration, we calculate the extreme rating degree denoted γ_i, corresponding to each reviewer R_i, which is considered as the discounting factor. Then, we weakened an other time the reviewer helpfulness ${}^{\lambda_i}m^{\Omega_S}_{RH_i}$ by its relative reliability degree using the discounting operation. Thus, the obtained discounted ${}^{\lambda_i\gamma_i}m^{\Omega_S}_{RH_i}$ modeled the reviewer helpfulness based on both the reviewer helpfulness degree and extreme rating which are an important spammers' indicators.

3.2.3 Combining the both reviewer reputation and helpfulness

With the purpose of representing the whole reviewer trustworthiness. We combine the reviewer *bba*'s reputation ${}^{\delta_i}m^{\Omega_S}_{RR_i}$ with his helpfulness *bba* ${}^{\lambda_i\gamma_i}m^{\Omega_S}_{RH_i}$ using the Dempster rule of combination under the frame of discernment Ω_S. The joint resultant *bba* $m^{\Omega_S}_{RT_i}$ illustrates each reviewer's the trustworthiness degree.

3.3 Distinguishing between the fake and the genuine reviews

As highlighter before, relying on both spam review and spammer review detection become the most effective solution to spot deceptive reviews. For this reason, we combine both the reviewer's opinion trustworthiness modeled by m^{Θ}_i with the reviewer spamicity represented by $m^{\Omega_S}_{RT_i}$ in order to make a powerful decision. For doing so, we apply the following steps.

3.3.1 Modeling both the reviewer and his giving opinion trustworthiness

In the interest of modeling both the reviewer and his opinion trustworthiness in one joint *bba*, we deal with the following steps:
First of all, we define the global frame of discernment relative to the reviewer and his opinion trustworthiness. It represents the cross product of the two different frame Θ and Ω_S denoted by: $\Omega_{RR} = \Theta \times \Omega_S$. Then, we extend all the

mass functions m_i^{Θ} and $m_{RS_i}^{\Omega_S}$ to the global frame of discernment Ω_{RR} using the vacuous extension in order to get new *bba's* $m_i^{\Theta\uparrow\Omega_{RR}}$ and $m_{RS_i}^{\Omega_S\uparrow\Omega_{RR}}$. Finally, we combine these extended *bbas* using the Dempster combination rule $m_i^{\Omega_{RR}} = m_i^{\Theta\uparrow\Omega_{RR}} \oplus m_{RS_i}^{\Omega_S\uparrow\Omega_{RR}}$ to get the joint *bba* $m_i^{\Omega_{RR}}$ that represents both the reviewer and his given opinion trustworthiness.

3.3.2 Reviewer and his opinion trustworthiness transfer

In following step, we transfer the $m_i^{\Omega_{RR}}$ under the product space Ω_{RR} to the frame of discernment $\Theta_D = \{f, \bar{f}\}$ to make decision by modeling the reviewer opinion into fake or not fake.

In spam reviews detection field, all the reviews given by the spammers are considered as fake opinion reviews, because spammers are not real consumers and they try sometimes to post reviews compatible with the provided one to avoid being detected by the spam detection methods.

For that, a multi-valued operation, denoted τ is applied. The function τ: Ω_{RR} to 2^{Θ_D} rounds up event pairs as follows:

- Masses of pairs that contain at least an element $\{S\}$ spammer are transferred to fake $f \subseteq \Theta_D$ as:
  $m_\tau(\{f\}) = \sum_{\tau(SR_i)=f} m_i^{\Omega_{RR}}(SR_i),\ (SR_i = A \times S) \subseteq \Omega_{RR}$
- Masses of pairs including at least an element $\{\bar{S}\}$ not spammer are transferred to not fake $\bar{f} \subseteq \Theta_D$ such that:
  $m_\tau(\{\bar{f}\}) = \sum_{\tau(SR_i)=\bar{f}} m_i^{\Omega_{RR}}(SR_i),\ (SR_i = A \times \bar{S}) \subseteq \Omega_{RR}$
- Masses of event couples with no element in $\{S, \bar{S}\}$ are transferred to Θ_D as:
  $m_\tau(\Theta_D) = \sum_{\tau(SR_i)=\Theta_D} m_i^{\Omega_{RR}}(SR_i),\ (SR_i = A \times \Omega_S) \subseteq \Omega_{RR}$

3.3.3 Decision making

Finally, we apply the pignistic probability $BetP$ in order to distinguish between the fake and the genuine opinion. Hence, the $BetP$ with the grater value will be considered as the final decision.

4 Experimentation and Results

4.1 Experimentation tools

In our method, we conduct two real datasets collected and used in [14, 15] from yelp.com. These datasets are considered as the largest, richest, complete and only labeled datasets in the spam review research. These datasets offer the near-ground-truth since they are labeled through the yelp filter classifier, which has been used in many previous works [8, 14, 15] as a ground truth thanks to its efficient detection method based on several behavioral features, where recommended (Not filtered) reviews correspond to genuine reviews, and not recommended (filtered) reviews correspond to fake ones. Due to the huge number of reviews, we random sample the two datasets with 10% from the total number of reviews given to three different evaluation criteria (services, cleanliness and food quality). Table 1 introduces the datasets and indicates the ratio of (filtered) fake reviews (and consequently reviewers). Furthermore, we evaluate our method through the three following criteria: accuracy, precision and recall.

Table 1. Datasets description

Datasets	Reviews (filtered %)	Reviewers (Spammer %)	Services (Restaurant or hotel)
YelpZip	608,598 (13.22%)	260,277 (23.91%)	5,044
YelpNYC	359,052 (10.27%)	160,225 (17.79%)	923

4.2 Experimental results

As our method proposes a specific classifier able to differentiate between fake and genuine reviews given to overall or multiple evaluation criteria under an uncertain context. We propose to compare it with the state-of-art baselines classifiers; the Support Vector Machine (SVM) and the Naive Bayes (NB) used by most of the spam detection methods [7, 11, 13, 14]. In order to maintain safe comparison when applying the SVM and NB classifiers, we construct a balancing data (50% of fake reviews and 50% of genuine ones) extracted from our datasets (YelpZip and YelpNYC) to avoid the over-fitting, then we divided into 30% of testing set and 70% of training set and we use the features considered in our proposed method; the rating deviation, the reviewers average proliferation, the brust spamicity degree, the reviews helpfulness and the extreme rating providing by each reviewer. In addition, the final estimation of each evaluation criterion is obtained by averaging ten trials values using 10-Fold cross validation technique. Furthermore, we compare our method with the proposed uncertain classifier Multiple Criteria Belief Fake Reviews Detection (MC-BFRD) [2] which relies only on the review rating information. The results are reported in the Table 2.
Our method attends the highest performance detection according to accuracy,

Table 2. Comparative results

Evaluation Criteria	Accuracy				Precision				Recall			
Methods	NB	SVM	MC BFRD	Our method	NB	SVM	MC BFRD	Our method	NB	SVM	MC BFRD	Our method
YelpZip	64%	78%	70%	**91.5%**	62%	80%	74%	**95%**	64%	75%	72.78%	**90.1%**
YelpNYC	65%	82.5%	75%	**92.77%**	65%	86%	83%	**96%**	70%	84%	80%	**91.2%**

precision and recall over-passing the baseline classifier. It reaches an accuracy improvement until 14% with yelpZip and until 10% with yelpNYC data-set compared to SVM. Moreover, the improvement records between the two uncertain classifier (over 20%) confirms the importance of combining both the review and the reviewer features while considering the spammers' indicators in this field. Despite the fact that our approach is based on fewer indicators than yelp's filter classifier, we obtain competitive results (over 92%) thanks to our method ability in handling uncertainty within the different inputs. These encouraging results push us to integrate more behavioral features in our future work that we could

improve our results and obtain identical or even better performance than yelp filter.

5 Conclusion

The spam review is an actual big issue threaten the online reviews. To tackle this problem, we have propose a specific classifier able to deal with uncertainty in the multi-criteria rating reviews and in the reviewer information while analyzing them through the spammers indicators. Our proposed method shows performance in classifying the fake and the innocent reviews while testing with two real data-sets form yelp.com.

References

1. Ben Khalifa, M., Elouedi, Z., Lefèvre, E.: Fake reviews detection under belief function framework. Proceedings of AISI, 395-404 (2018)
2. Ben Khalifa, M., Elouedi, Z., Lefèvre, E.: Multiple criteria fake reviews detection using belief function theory. Proceedings of ISDA, 315-324 (2018)
3. Ben Khalifa, M., Elouedi, Z., Lefèvre, E.: Spammers detection based on reviewers' behaviors under belief function theory. Proceedings of IEA/AIE, 642-653 (2019)
4. Dempster, A.P.: Upper and lower probabilities induced by a multivalued mapping. Ann. Math. Stat.38, 325-339 (1967)
5. Deng, X., Chen, R.: Sentiment analysis based online restaurants fake reviews hype detection. Web Technologies and Applications, 1-10 (2014)
6. Lefèvre, E., Elouedi, Z.: How to preserve the confict as an alarm in the combination of belief functions? Decis. Support Syst.56, 326-333 (2013)
7. Fei, G., Mukherjee, A., Liu, B., Hsu, M., Castellanos, M., Ghosh, R.: Exploiting burstiness in reviews for review spammer detection. Proceedings of ICWSM, 13, 175-184 (2013)
8. Fontanarava, J., Pasi, G., Viviani, M.: Feature Analysis for Fake Review Detection through Supervised Classification. Proceedings of DSAA, 658-666 (2017).
9. Jindal, N., Liu, B.: Opinion spam and analysis. Proceedings of ACM, pp. 219-230 (2008).
10. Jousselme, A.-L., Grenier, D., Bossé, É.: A new distance between two bodies of evidence. Inf. Fusion 2(2), 91-101 (2001)
11. Lim, P., Nguyen, V., Jindal, N., Liu, B., Lauw, H. : Detecting product review spammers using rating behaviors. Proceedings of CIKM, 939-948 (2010)
12. Ling, X., Rudd, W.: Combining opinions from several experts. Applied Artificial Intelligence an International Journal, 3 (4), 439-452 (1989)
13. Mukherjee, A., Kumar, A., Liu, B., Wang, J., Hsu, M., Castellanos, M.: Spotting opinion spammers using behavioral footprints. Proceedings of ACM SIGKDD, 632-640 (2013)
14. Mukherjee, A., Venkataraman, V., Liu, B., Glance, N.: What Yelp Fake Review Filter Might Be Doing. Proceedings of ICWSM, 409-418 (2013)
15. Rayana, S., Akoglu, L.: Collective opinion spam detection: Bridging review networks and metadata. Proceedings of ACM SIGKDD, 985-994 (2015)
16. Shafer, G.: A Mathematical Theory of Evidence, vol. 1. Princeton University Press (1976)
17. Smets, P.: The transferable belief model for quantified belief representation. In: Quantified Representation of Uncertainty and Imprecision, 267-301. Springer, Dordrecht (1998)
18. Wang, G., Xie, S., Liu, B., Yu, P. S.: Review graph based online store review spammer detection. Proceedings of ICDM, 1242-1247 (2011)