Trolls Identification within an Uncertain Framework

Imen Ouled Dlala, Dorra Attiaoui, Arnaud Martin, Boutheina Ben Yaghlane

Trolls Identification within an Uncertain Framework

Imen Ouled Dlala, Dorra Attiaoui, Arnaud Martin, and Boutheina Ben Yaghlane

Abstract—The web plays an important role in people's social lives since the emergence of Web 2.0. It facilitates the interaction between users, gives them the possibility to freely interact, share and collaborate through social networks, online communities forums, blogs, wikis and other online collaborative media. However, an other side of the web is negatively taken such as posting inflammatory messages. Thus, when dealing with the online communities forums, the managers seek to always enhance the performance of such platforms. In fact, to keep the serenity and prohibit the disturbance of the normal atmosphere, managers always try to novice users against these malicious persons by posting such message (DO NOT FEED TROLLS). But, this kind of warning is not enough to reduce this phenomenon. In this context we propose a new approach for detecting malicious people also called 'Trolls' in order to allow community managers to take their ability to post online. To be more realistic, our proposal is defined within an uncertain framework. Based on the assumption consisting on the trolls' integration in the successful discussion threads, we try to detect the presence of such malicious users. Indeed, this method is based on a conflict measure of the belief function theory applied between the different messages of the thread. In order to show the feasibility and the result of our approach, we test it in different simulated data.

Keywords—Q&AC, trolls, belief function theory, conflict measure.

I. INTRODUCTION

THE way we look for, and acquire information has shifted greatly into to instant, easy and low cost process. In fact, thanks to the Internet one can make a research in any given topic, get a huge amount of information by a simple click. Although, for some problems it is difficult to get satisfactory answers by searching directly on a traditional search engine. Instead, we prefer to find someone who has expertise or experience. In order to have the best answer, one of the tools that has widened the scope of information exchange is Question Answering Communities (Q&AC). These systems allow everyone to contribute as much as they can on a given community. Unfortunately, not all messages can be considered as reliable: some users claim themselves as experts, and other people post messages without any utility for the one who is seeking for answers. Thus, the managers of these communities seek to always enhance the performance of such platforms. Although, the increase of the useless messages can be attributed to the presence of trolls. The term of trolling has been defined in several works within different communities, including [2], [6] and [17]. These Malicious people intend to insidiously mislead the subject of the discussion in order to provoke controversy and disrupt the discussion. They aim to make normal users fall into their traps by deviating them from the main topic of the discussion. In fact, the only way to deal with a troll is to ignore him, or detect his presence in order to notify him or take away his ability to post online. Thus,

some other works tried to detect not just their characteristics but also their presence in order to avoid them. To address this problem Cambria et al. [14] proposed a technique based on semantic computing to automatically detect and check web trolls. This work aims to prevent the malicious people from emotionally hurting other users or communities within the same social network. In another work Ortega et al. [15] proposed a method to classify users in a social network regarding to their trustworthiness. The goal of their method is to detect trolls from the other users by preventing such malicious users to gain high reputation in the network. Patxi et al. [16] dealt with Trolling users on twitter social network. These studies were explored in different social networks within certain framework.

When dealing with real-world applications, the massive amounts of data are inseparably connected with imperfection. In fact, this kind of data can be imprecise and/or uncertain or even missing. Different theories have emerged to deal with this kind of data such as fuzzy set theory [21], possibility theory [22] and belief function theory [1]. Thus, to be closer to reality and to obtain more relevant results, we propose a new method dealing with uncertain data. This method aims to detect trolls in Q&AC using the framework of belief function.

The paper proceeds as follows: in Section 2, we introduce the Q&AC and briefly review related works. In section 3, we present the necessary background regarding the different concepts of the belief functions theory. We define the different steps of our proposal based on a conflict measure in section 4. Finally, we present the feasibility of the proposed method on an illustrated example.

II. Q&AC: QUICK OVERVIEW

In this section we introduce some concepts related the Q&AC. First we will start by presenting the main actors in these forums, then a little overview on sources identification and finally the levels of uncertainty we can face in Q&AC.

A. Users within Q&AC

Users are considered as the main actors within Question Answering Communities. We can define different types such as experts, trolls and learners.

- **Reliable user / Expert**: a person who is very knowledgeable about or skilful in a particular area.
- **Troll**: a person who seeks to disturb the serenity of the concerned community. His purpose is to create controversial debates by multiplying irrelevant messages that we keep unanswered.
- **Learner**: a normal user of the Question Answering Community, trying to gain information and expertise.

B. Sources Identification within Q&AC

Several researches have been exploring this field, trying to evaluate sources of information in Q&AC. Such as Bouguessa et al. [7] who proposed a model to identify authoritative users based on the number of best answers provided by them. A best answer is selected either by the asker or by other users via a voting procedure. In [12], the author focused on the selection of questions a user would choose for answering. Based on these studies, experts prefer answering questions where they have a higher chance of making a valuable contribution. Recently in [13], the authors proposed a framework for evaluating both the reliability and the expertise of an information provider. Considering some cognitive and behavioral criteria of the users, they were able to establish a trust system. Using a response matrix summarizing the interactions between peers of persons, each one is capable of estimating and providing an opinion. Using the subjective logic to aggregate these evaluations, they provided later a global reliability and expertise value for each user within Q&AC.

C. Uncertainty within Q&AC

When dealing with information provided by humans, we are facing several levels of uncertainty. Gjergji et al. proposed three levels for Q&AC [11], the first one is related to the extraction and integration of uncertainty, the second deals with information sources uncertainty and finally the inherent knowledge related to the information itself. In our case, we are more interested in the evaluation of the sources and the part of uncertainty related to them. The main issue in these communities is that we are facing users that we do not always have an *apriori* knowledge about them. We ignore every thing about the sources' credibility, reliability, relevance, objectivity and expertise. In this context, we will exploit all the mathematical background and large panel of sepcificities provided by the theory of belief functions to help us considering this problem in an uncertain point of view.

III. THEORY OF BELIEF FUNCTIONS

This section recalls the necessary background related to the belief function theory It has been developed by Dempster in his work on upper and lower probabilities [1]. Based on that, he was able to represent more precisely the observed data.

A belief function must take into consideration all the possible events on which a source can describe a belief. Based on that, we can define the frame of discernment.

A. Frame of discernment

It is a finite set of disjoint elements noted Ω where $\Omega = \{\omega_1, ..., \omega_n\}$. This theory allows us to affect a mass on a set of hypotheses not only a singleton like in the probabilistic theory. Thus, we are able to represent ignorance, imprecision...

B. Basic belief assignment (bba)

A bba is defined on the set of all subsets of Ω, named power set and noted 2^Ω. It affects a real value from $[0, 1]$ to every subset of 2^Ω reflecting sources amount of belief on this subset. A bba m verifies:

$$\sum_{X \subseteq \Omega} m(X) = 1. \tag{1}$$

We consider any positive elementary mass $m(X) > 0$ as a focal element such that X belongs to 2^Ω.

C. Combination rules

Many combination rules have been proposed taking in consideration the nature of the sources.

1) Dempster's combination rule: The first one was proposed by Dempster in 1967 [1] which is a conjunctive normalized combination rule also called the orthogonal sum. Given two mass functions m_1 and m_2, for all $X \in 2^\Omega$, $X \neq \emptyset$, the Dempster's rule is defined by:

$$m_D(X) = m_1 \cap m_2(X) = \frac{1}{1-k} \sum_{Y_1 \cap Y_2 = X} m_1(Y_1)m_2(Y_2) \tag{2}$$

where $k = \sum_{Y_1 \cap Y_2 = \emptyset} m_1(Y_1)m_2(Y_2)$ is the inconsistency of the fusion (or of the combination) can also be called the conflict or global conflict. $(1 - k)$ is the normalization factor of the combination in a closed world.

2) The conjunctive combination rule: In order to consider the issues of the open world, the conjunctive combination rule was introduced by Smets [9]. Considering two mass functions m_1 and m_2, for all $X \in 2^\Omega$ m_{conj} is defined by:

$$m_{conj}(X) = \sum_{Y_1 \cap Y_2 = X} m_1(Y_1)m_2(Y_2) \tag{3}$$

3) The disjunctive combination rule: First introduced by Dubois and Prade 1986 [18], the induced results of two bbas m_1 and m_2 is defined as follows:

$$\forall X \subseteq \Omega \, m_{disj}(X) = \sum_{Y_1 \cup Y_2 = X} m_1(Y_1)m_2(Y_2) \tag{4}$$

The disjunctive combination rule can be used when one of the sources is reliable or when we have no knowledge about their reliability.

IV. INCLUSION AS A CONFLICT MEASURE FOR BELIEF FUNCTIONS

Recently Martin in [3] used a degree of inclusion as involved in the measurement made in order to determine the conflict during the combination of two belief functions. He presented an index of inclusion having binary values where:

$$Inc(X_1, Y_2) = \begin{cases} 1, \text{if} X_1 \subseteq Y_2 \\ 0, \text{otherwise} \end{cases} \tag{5}$$

With X_1, Y_2 being respectively the focal elements of m_1 and m_2. This index is then used to measure the degree of inclusion of the two mass functions and defined as:

$$d_{inc} = \frac{1}{|F_1||F_2|} \sum_{X_1 \in F_1} \sum_{Y_2 \in F_2} Inc(X_1, Y_2) \qquad (6)$$

Where $|F_1|$ and $|F_2|$ are the number of focal elements of m_1 and m_2. He define the degree of inclusion of m_1 and m_2: $\sigma_{inc}(m_1, m_2)$ as follows:

$$\sigma_{inc}(m_1, m_2) = max(d_{inc}(m_1, m_2), d_{inc}(m_2, m_1)) \qquad (7)$$

Where d_{inc} is the degree of inclusion of m_1 in m_2 and inversely. This inclusion is used as a conflict measure for two mass functions, using it like presented:

$$Conf(m_1, m_2) = (1 - \sigma_{inc}(m_1, m_2)d(m_1, m_2)) \qquad (8)$$

where $d(m_1, m_2)$, is the distance of Jousselme [10]:

$$d(m_1, m_2) = \sqrt{\frac{1}{2}(m_1 - m_2)^T \underline{\underline{D}}(m_1 - m_2)} \qquad (9)$$

where $\underline{\underline{D}}$ is a metric based on the measure of Jaccard:

$$D(A, B) = \begin{cases} 1, if A = B = \emptyset \\ \frac{|A \cap B|}{|A \cup B|}, \forall A, B \in 2^\Omega \end{cases} \qquad (10)$$

V. TROLLS IDENTIFICATION BASED IN A CONFLICT MEASURE

Based on the assumption that consists of the trolls' integration in the successful discussion threads, we propose a new method for detecting malicious people in online communities forums. This approach is defined within the framework of belief functions. Indeed, it is based on a conflict measure of this theory applied between the different messages of the thread. We can summarize our proposed method in three major steps that will be discussed in depth in the following.

A. Users' messages

Hardarker proposed primary characteristics of a troll [2] (Aggression, Deception, Disruption, Success). In 2014, Buckels et al. [6] specified the behavioral characteristics of a troll. They described them as persons having sadism, psychopathy and machiavilism. To them, trolling is a "deceptive, destructive or disruptive manner in social media".
In the context of this work, to distinguish between the troll and the other users, we tried to manually extracted the characteristics of their responses from the answers and comments in different forums. Based on these characteristics, the content of the messages can be: Off-topic, senseless or controversy. Using these characteristics, we have defined the frame of discernment that can characterize a message in a forum:

$$\Omega_{msg} = \{Off - topic, Senseless, 1, \ldots, N\} \qquad (11)$$

- Senseless: how much the response is empty of meaning?
- Off-topic: How irrelative the answer can be?
- $[1..N]$: number of topics where, $[1..N] \backslash i$ with i being the relevant topic, and $[1..N] \backslash i$ are the controversy topics posted by a troll.

During this step, we assume that a method of analysis expresses a piece of evidence concerning the nature of each message. This method aims to analyze the messages relative to the posted question or topic.

B. Users' conflict

Detecting irrelevant messages does not only means that this user is a troll. Thus, it is not only the content of the messages that can characterize the trolls. We can find a victim user that responds to a message posted by a troll. Besides, the subject of the discussion can change gradually. In fact, to distinguish between trolls and other users in a community, we need to quantify how a given user is in conflict with the rest of all the other users. Thus, we will base our approach on measuring the conflict between the messages of each person posting answers. The list of notations is shown in table I. Using the inclusion

TABLE I
LIST OF NOTATIONS

Notations	Description
U	Users
N	Number of users
NP	Number of all the previous messages
NP_j	Number of the previous messages of a user U_j
N_i	Number of all messages posted by a user U_i
N_j	Number of all messages posted by a user U_j
m_k	k^{th} message of a user U_i
m_s	s^{th} message of a user U_j
$Rank(m)$	Rank of the message m
$Tab1$	Contains in each time the conflict of a message relative to each user
$Tab2$	Contains in each time the number of the previous messages of a message
$Tab3$	Contains the total conflict of each user
$Conf_t$	Contains the sum of conflict of each user

as a conflict measure for belief functions, for each user U_i we will measure:

- $Conf_{msg/U}$: measures the conflict between the kth message posted by U_i and the messages that were posted before it by each other users U_j.

$$Conf_{msg/U}(m_k(U_i), m(U_j)) =$$
$$\frac{1}{NP_j} \sum_{s=1}^{NP_j} Conf(m_k(U_i), m_s(U_j)), (i \neq j) \qquad (12)$$

- $Conf_{msg}$: measures the conflict between the kth message posted by U_i and the all messages that were posted before it by all the other users U based on a weighted mean. This measure takes into account the number of messages posted by every user in order to determine the level of conflict especially between a troll and an expert.

$$Conf_{msg}(m_k(U_i), m(U)) =$$
$$\sum_{j=1}^{N} \frac{NP_j}{NP} Conf_{msg/U}(m_k(U_i), m(U_j)) \qquad (13)$$

- $Conf_{user}$: measures the global conflict of the user U_i

$$Conf_{user} = \frac{1}{N_i} \sum_{k=1}^{N_i} Conf_{msg}(m_k(U_i), m(U)) \qquad (14)$$

The value of the total conflict of a user can be risen when this user launches into an interminable debate with a troll. In this case, this victim user becomes in his turn a troll. Thus, the managers have to control the behavior of the users in many discussion threads.

C. Users' clustering

The last step consists on the classification of the users according to their conflict results into two groups. Therefore, to make decision we base our approach on an unsupervised classification method using the k means algorithm.

It was introduced by McQueen [19] and implemented in its current forms by Forgy [20]. The Kmeans algorithm aims to construct from the objects of the training set K partitions (clusters) concentrated and isolated from each other. In our case, we will devise the users into two partitions: K= 2. Since the value of the troll 's conflict is bigger than the conflict of the other users:

- Trolls belong to the group having the biggest value of center.

- The other users belong to the group having the least value of center.

VI. EXPERIMENTATION

To illustrate the comportment of our proposed method, we have tested it in different simulated data. In this section, we will present two different examples.

A. Example 1

As we presented our method, it has three main steps. Indeed, we will present the results of each step:

1) Users' messages: Our assumption consists on the integration of the trolls in the successful discussion threads. From this point of view, we simulate the data of analyze of messages as depicted in Figure 1. In fact, in this example we will try to detect a troll in a group of 4 users. In this scenario, the discussion thread contains 16 messages posted by different users and among whom three messages are published by a troll. In this example, our frame of discernment is composed by 4 elements: Relevant=X_1, off-topic=X_2, senseless=X_3, controversy-topic=X_4. As shown in Figure 2 each row presents: the owner of the message, the order of the message in the discussion thread and the mass function of this message (as mention in section III. B each bba must be equal to 1).

In this example, the first message of the troll (U_4) is controversy: $m(X_4) = 0.9210$. His second message is empty of meaning: $m(X_3) = 0.9716$. His third message is controversy: $m(X_4) = 0.8387$.

2) Users conflict: Based on the method of the inclusion and applying our algorithm, we will present the total conflict of each user of our example in Table II: U_4 has the biggest value of conflict. The total conflicts of users U_1 and U_2 is small relative to the total conflict of U_4 despite the fact that they responded to the first message of the troll by posting each one a controversy message. this result can be explained by the answers provided by these two users who have published relevant messages. U_3 has a small value of conflict, he published three relevant messages where in his first message $m(X_1) = 0.9732$, in his second message $m(X_1) = 0.7782$, and his third message $m(X_1) = 0.9632$.

Users	Ranks	ø	X_1	X_2	$X_1 \cup X_2$	X_3	$X_1 \cup X_3$	$X_2 \cup X_3$	$X_1 \cup X_2 \cup X_3$
1	2	0	0.9872	0	0	0	0	0	0
1	6	0	0.7734	0	0	0.0817	0	0	0
1	9	0	0	0	0.0152	0	0	0	0
1	11	0	0.7476	0	0	0	0	0	0
1	15	0	0.6881	0.0261	0.0242	0	0	0	0
2	1	0	0.8853	0	0	0	0	0	0
2	3	0	0.9124	0	0	0	0	0	0
2	5	0	0.7714	0	0.0588	0	0	0	0
2	8	0	0	0	0	0	0	0	0.0560
2	14	0	0.6942	0.0721	0	0	0	0	0
3	4	0	0.9732	0	0	0.0015	0.0162	0	0
3	10	0	0.7782	0	0	0	0	0	0
3	13	0	0.9632	0	0	0	0.0146	0	0
4	7	0	0	0	0	0	0	0	0.0190
4	12	0	0	0.0213	0	0.9716	0	0	0
4	16	0	0	0	0	0	0	0	0

Users	Ranks	X_4	$X_1 \cup X_4$	$X_2 \cup X_4$	$X_1 \cup X_2 \cup X_4$	$X_3 \cup X_4$	$X_1 \cup X_3 \cup X_4$	$X_2 \cup X_3 \cup X_4$	$X_1 \cup X_2 \cup X_3 \cup X_4$
1	2	0	0	0	0.0030	0	0	0	0.0098
1	6	0	0	0	0.0581	0	0	0	0.0868
1	9	0.9400	0	0	0	0.0261	0	0	0.0188
1	11	0	0.1036	0	0	0	0.0467	0	0.1022
1	15	0	0	0	0	0	0	0	0.2616
2	1	0.0569	0	0	0	0	0.0164	0	0.0414
2	3	0.0204	0	0	0.0182	0	0	0	0.0490
2	5	0	0	0.1590	0	0	0	0	0.0108
2	8	0.8523	0	0.0467	0	0	0	0	0.0450
2	14	0	0	0	0	0.0982	0	0	0.1355
3	4	0	0	0	0	0	0	0	0.0091
3	10	0	0.0548	0	0	0	0.0650	0	0.1019
3	13	0	0	0	0.0105	0	0	0	0.0118
4	7	0.9210	0	0	0	0	0	0	0.0600
4	12	0	0	0	0	0	0	0	0.0071
4	16	0.8387	0.0606	0	0	0	0	0	0.1007

Fig. 1. Simulation Results

TABLE II
TOTAL CONFLICT OF EACH USER

	U_1	U_2	U_3	U_4
Conf$_{user}$	0.0610	0.0639	0.0489	0.2030

3) Users clustering: Applying the K-means algorithm to the different values of total conflict of all users we obtained two clusters.

- Trolls= $\{U_4\}$
- Other users= $\{U_1, U_2, U_3\}$

Our proposal provides us a correct classification of the users. This result shows the feasibility of our proposed method.

B. Example 2

For this simulation we will assume that we are dealing with 8 users, among them two trolls. The discussion thread contains 31 messages. The result of the total conflict of each user expressed in equation 14 is illustrated in figure 2.

The first troll U_4 published 2 controversy messages and the second troll U_8 published 3 messages: The two first ones are off-topic, and the last one is controversy.

- U_1 posted 3 relevant messages and 2 controversy messages to respond to the first troll.
- U_2 posted 7 relevant messages and 2 controversy messages to respond to the first troll.
- U_3 posted 4 relevant messages and one off-topic message to respond to the second troll.
- U_5 posted one relevant message.
- U_6 published 3 relevant messages.
- U_7 published 2 relevant messages.

The total conflict of the troll U_4 is bigger relatively to the other troll U_8 because he published his posts after a big number of reliable messages provided by the other users. So, this situation created a higher value of a conflict. Applying the Kmeans algorithm our method provides us a correct classification:

- Trolls= $\{U_4, U_8\}$
- Other users= $\{U_1, U_2, U_3, U_5, U_6, U_7\}$

The users U_1, U_2 and U_3 are not classified among the trolls in spite of their posts that can be categorized as trolls' messages. This result is explained by the fact that they have other relevant messages.
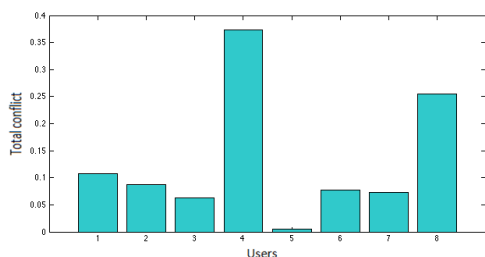

Fig. 2.   The steps of detecting trolls

VII. CONCLUSION

We proposed in this paper a new method for detecting 'Trolls' in Q&AC. Relying on this approach managers, can control the behavior of the users in many discussion threads in order to notify them to stop trolling. Our work is defined within an uncertain framework. It is based on a conflict measure in the belief function theory applied between the messages of the different users of the thread. First of all, this method aims to analyze the messages relative to the posted question or topic. But detecting irrelevant message is not enough to judge if this user is a troll or not. Thus, not only the content of the messages that can characterize the trolls but also their behaviors. Next, using the results of this analysis we measured the conflict between the different users. Finally, after calculating the conflict of each user we applied the kmeans method in order to distinguish trolls from the other users. Indeed, we have classified the users according to their conflict results into two clusters. This method was tested in different simulated data to check its feasibility. Since our proposed method for detecting malicious users dealt only with one discussion thread, we aim to extend this approach to detect trolls inside the community.

REFERENCES

[1] A. P. Dempster, *Upper and Lower probabilities induced by a multivalued mapping*, Annals of Mathematical Statistics, vol. 38, pp. 325-339, 1967.
[2] C. Hardarker, *Trolling is asynchronous computer-mediated communication: from user discussion to academic definitions*, Politeness Research, vol. 6, pp. 215-242, 2010.
[3] A. Martin, *About Conflict in the Theory of Belief Functions*, 2nd International Conference on Belief Functions, Advances in Soft Computing, vol. 164, pp. 161-168, 2012.
G. Shafer, *A mathematical theory of evidence*, Princeton University Press, 1976.
[4] P. Smets, *Constructing the Pignistic Probability Function in a Context of Uncertainty*, Fifth Annual Conference on Uncertainty in Artificial Intelligence, pp. 29-40, 1989.
[5] E. Buckels, D. Tropnell and D. Paulhus, *Trolls just want to have fun*, Personality and individual Differences, 2014.
[6] M. Bouguessa, B. Dumoulin, D. Paulhus, and S. Shengrui, *Identifying authoritative actors in question answering forums: the case of Yahoo! answers*, ACM International Conference on Knowledge Discovery and Data mining, KDD, pp. 866-874, 2008.
[7] P. Smets and R. Kennes, *The Transferable Belief Model*, Artificial Intelligence, vol. 66, pp. 191-234, 1994.
[8] A. L. Jousselme, D. Grenier and E. Bosse, *A new distance between two bodies of evidence*, Information Fusion, vol. 66, pp. 91-101, 2001.
[9] K. Gjergji, V.G Jurgen, D. Stern and G. Thore, *CoBayes: bayesian knowledge corroboration with assessors of unknown areas of expertise*, WSDM, vol. 2, pp. 465-474, 2011.
[10] A. Pal and J. A Konstan, *Expert identification in community question answering: exploring question selection bias*, ACM, pp. 1505-1508, 2010.
[11] K. Pelechrinisn, V. Zadorozhny, V. Kounev, V. Oleshchuk, M. Anwar, and Y. Lin, *Automatic Evaluation of Information Provider Reliability and Expertise*, World Wide Web, Springer, vol. 2, 2013.
[12] E. Cambria, P. Chandra, A. Sharma and A. Hussain, *Do not feel the trolls*, Proceedings of the 9th International Semantic Web Conference, 2010.
[13] F. Ortega, A. Jose, L. Fermin, G. Carlos and E. Fernando, *Propagation of trust and distrust for the detection of trolls in a social network*, Computer Networks, vol. 56, pp. 2884-2895, 2012.
[14] G. Patxi, G. Jose, L. Carlos, S. Igor and G. Pablo, *Supervised machine learning for the detection of troll profiles in twitter social network: Application to a real case of cyberbullying.*, International Joint Conference SOCO-CISIS-ICEUTE, Springer, pp. 419-428, 2013.
[15] P. Shachaf and N. Hara, *Beyond vandalism: Wikipedia trolls*, Information Science, vol. 36, pp. 357-370, 2010.
[16] D. Dubois and H. Prade, *A set theoretic view of belief functions: logical operation and approximations by sets*, International Journal of General Systems, vol. 12, pp. 193-226, 1986.
[17] J. McQueen, *Some methods for classification and analysis of multivariate observations*, The Fifth Berkeley Symposium on Mathematical Statistics and Probability, vol. 1 , pp. 281-297, 1967.
[18] J. McQueen, *Cluster analysis of multivariate data : Efficiency vs. Interpretability of Classifications*, Biometrics, vol. 1 , pp. 21-68, 1965.
[19] L. Zadeh, *Fuzzy sets*, Information and Control, vol. 8 , pp. 338-353, 1965.
[20] D. Dubois and H. Prade, *Possibility Theory*, Springer, 1988.

Imen Ouled Dlala Univeristy of Tunis, LARODEC laboratory, High Institut of Management, Cite Bouchoucha, le Bardo, TUNISIA; E-mail: dlalaimen@yahoo.fr

Dorra Attiaoui Univeristy of Tunis, LARODEC laboratory, High Institut of Management, Cite Bouchoucha, le Bardo, TUNISIA; University of Rennes 1, IRISA, Rue E. Branly, 22300 Lannion, FRANCE; E-mail: attiaoui.dorra@gmail.com

Arnaud Martin University of Rennes 1, IRISA, Rue E. Branly, 22300 Lannion, FRANCE; E-mail: Arnaud.Martin@univ-rennes1.fr

Boutheina Ben Yaghlane University of Carthage, IHEC Carthage, Carthage Presidence, TUNISIA, LARODEC laboratory; E-mail: boutheina.yaghlane@ihec.rnu.tn