Belief Rough Set Classification for Web Mining based on Dynamic Core

Salsabil Trabelsi, Zied Elouedi
Larodec, Institut Superieur de Gestion de Tunis, Tunisia
E-mail: salsabilt@yahoo.fr, zied.elouedi@gmx.fr

Pawan Lingras
Department of Mathematics and Computing Science, Saint Mary's University
Halifax, Nova Scotia, B3H 3C3, Canada.
E-mail: pawan@cs.smu.ca

***Abstract*—In this paper, we apply classification system denoted Belief Rough Set Classifier (BRSC) based on the hybridization of belief functions and rough sets to learn decision rules from uncertain data consisting of web usage. The uncertainty appears only in decision attributes and is handled by the Transferable Belief Model (TBM), one interpretation of the belief function theory. The web usage mining dataset was obtained from an educational website, where the visits were clustered based on study patterns. Instead of using crisp assignment of a visit to one of the clusters, the study associated a belief that the visit will belong to that cluster. Due to the uncertainty existing in the chosen web usage mining database, the feature selection step used to construct our BRSC will be based on dynamic core approach which allows getting better performance in uncertain and large database. To judge the performance of our classifier, we choose three evaluation criteria: accuracy, size and time requirement. Besides, we compare the results with those obtained from a similar classifier, namely Belief Decision Tree (BDT).**

I. Introduction

With the large amount of information available online, the World Wide Web is a fertile area for data mining research. Web mining is a topic of increasing research interest. We can categorize web data mining into three areas; web content mining, web structure mining and web usage mining. The latter has seen a rapid increase in interest which is the application of data mining techniques to discover usage patterns from web data, in order to understand and better serve the needs of web-based applications.

The Rough Sets (RS) [8], [9] constitutes a sound basis for decision support system and data mining. It is a successful technique of classification applied in several real-world applications like web usage area [6]. This technique of classification performs feature selection before generating rules. It is an efficient technique that tries to produce in automatic way a minimal and a significant set of decision rules without many iterations. However, the web usage databases like many other real-world databases may be characterized by incomplete or uncertain data. The standard classification based on rough sets is not able to discover patterns from this kind of databases. In this paper, we apply a new approach of classification denoted Belief Rough Set Classifier to discover hidden patterns from uncertain web usage database. The BRSC proposed originally in [16] was able to learn decision rules from uncertain data. The uncertainty appears only in decision attributes and is handled by the belief function theory as understood in the transferable belief model. In fact, the TBM offers a convenient framework thanks to its ability to represent epistemological uncertainty. It allows experts to express partial beliefs in a much more flexible way than probability functions do. It also allows partial or even total ignorance concerning classification parameters. In addition to these advantages, the strength of the TBM appears essentially in its ability to combine several pieces of evidence.

The web usage mining dataset used in this paper was obtained from web access logs of the introductory computing science course at Saint Mary's University. The visits were clustered based on study patterns. Instead of using crisp assignment of a visit to one of the three clusters (Studious, Crammers, Workers), the study associated a basic belief assignment (bba). The resulting uncertain clustering was characterized using belief functions. To better handle the noise and the uncertain issues existing in the chosen web usage mining database, the feature selection step used to construct our BRSC will be based on dynamic core which consists of the more stable attributes which cannot be reduced from the decision table and the more frequently appearing core in sub-decision tables created by random samples of a given decision table. Decision rules induced by means of dynamic core are more appropriate to classify new objects.

This paper is organized as follows: Section 2 provides an overview of the basic concepts of RS. Section 3 introduces the belief function theory as understood in the Transferable Belief Model. Section 4 describes our approach of classification denoted BRSC which is based on rough sets under the belief function framework to handle the problem of uncertainty. Finally, in Section 5, we describe the chosen web usage mining database and we report the experimental results obtained by applying the BRSC to this real and uncertain database.

978-1-4244-8136-1/10/$26.00

II. Rough set theory

In this Section, we recall some basic notions related to rough sets [9]. A decision table (DT) is defined as A= (U, C, {d}), where $U = \{o_1, o_2, ..., o_n\}$ is a nonempty finite set of n objects called $the\ universe$, $C = \{c_1, c_2, ..., c_k\}$ is a finite set of k $condition$ attributes and $d \notin C$ is a distinguished attribute called $decision$. The value set of d, called Θ =$\{d_1, d_2, ..., d_s\}$. The rough sets adopt the concept of indiscernibility relation [9] to partition the object set U into disjoint subsets, denoted by U/B or IND_B. The partition that includes o_j is denoted $[o_j]_B$.

$$IND_B = U/B = \{[o_j]_B | o_j \in U\} \tag{1}$$

Where

$$[o_j]_B = \{o_i | \forall c \in B\, c(o_i) = c(o_j)\} \tag{2}$$

The notation $c(o_j)$ is used to represent the value of a condition attribute $c \in C$ for an object $o_j \in U$. The equivalence classes based on the decision attribute are denoted by $U/\{d\}$.

$$IND_{\{d\}} = U/\{d\} = \{[o_j]_{\{d\}} | o_j \in U\} \tag{3}$$

Let $B \subseteq C$ and $X \subseteq U$. We can approximate X by constructing the $B{-}lower$ and $B{-}upper\ approximations$ of X, denoted $\underline{B}$(X) and $\bar{B}$(X), respectively, where

$$\underline{B}(X) = \{o_j | [o_j]_B \subseteq X\}\ and\ \bar{B}(X) = \{o_j | [o_j]_B \cap X \neq \emptyset\} \tag{4}$$

A reduct is a minimal subset of attributes from C that preserves the positive region and the ability to perform classifications as the entire attributes set C. A subset $B \subseteq C$ is a reduct of C with respect to d, iff B is minimal and:

$$Pos_B(A, \{d\}) = Pos_C(A, \{d\}) \tag{5}$$

Where $Pos_C(A, \{d\})$, called a positive region of the partition $U/\{d\}$ with respect to C.

$$Pos_C(A, \{d\}) = \bigcup_{X \in U/\{d\}} \underline{C}(X) \tag{6}$$

The core is the most important subset of attributes, it is included in every reduct.

$$Core_C(A, \{d\}) = \bigcap Red_C(A, \{d\}) \tag{7}$$

Where $Red_C(A, \{d\})$ is the set of all reducts of A relative to d.

III. Belief function theory

In this Section, we briefly review the main concepts underlying the belief function theory as interpreted in the TBM [11]. Let Θ be a finite set of elementary events to a given problem, called the frame of discernment. All the subsets of Θ belong to the power set of Θ, denoted by 2^Θ. The impact of a piece of evidence on the subsets of the frame of discernment Θ is represented by a basic belief assignment (bba). The bba is a function $m : 2^\Theta \to [0, 1]$ such that:

$$\sum_{E \subseteq \Theta} m(E) = 1 \tag{8}$$

The value *m(E)*, called a basic belief mass (bbm), represents the portion of belief committed exactly to the event E. The bba's induced from distinct pieces of evidence are combined by the rule of combination [11]:

$$(m_1 \circledcap m_2)(E) = \sum_{F,G \subseteq \Theta : F \cap G = E} m_1(F) \times m_2(G) \tag{9}$$

IV. Belief Rough Set Classifier based on Dynamic Core

In this Section, we present our approach of classification called Belief Rough Set Classifier [16]. To more overcome the problem of uncertainty, we suggest apply the approach of dynamic core for feature selection step.

A. Basic concepts of rough sets under uncertainty

This Section describes the modified definitions of decision table, indiscernibility relation (tolerance relation), set approximations, positive region and dependency of attributes under the belief function framework. These adaptations were originally proposed in [14].

1) Uncertain Decision Table (UDT): It is given by $A = (U, C \cup \{ud\})$, where $U = \{o_j : 1 \leq j \leq n\}$ is a set of objects characterized by a set of certain condition attributes C=$\{c_1,\ c_2$,...,$c_k\}$ and an uncertain decision attribute ud. The value set of ud, called Θ =$\{ud_1,\ ud_2$,...,$ud_s\}$. We represent the uncertainty of each object o_j by a bba m_j expressing beliefs on decisions defined on the power set of Θ.

Example: *Let us use Table I to describe our UDT. It contains eight objects, three certain condition attributes C=$\{Hair$, $Eyes$, $Height\}$ and an uncertain decision attribute ud with two possible values $\{ud_1,\ ud_2\}$ representing Θ. For the object o_3, 0.7 of beliefs are exactly committed to the decision ud_1, whereas 0.3 of beliefs is assigned to the whole of frame of discernment Θ (ignorance). With bba, we can represent the certain case, like for the objects o_2, o_5 and o_7.*

2) Tolerance relation: The indiscernibility relation for the condition attributes U/C is the same as in the certain case because their values are certain. However, the indiscernibility relation for the decision attribute $U/\{ud\}$ is not the same as in the certain case. The decision value is represented by a bba. For this reason, the indiscernibility relation will be denoted $tolerance\ relation$. To redefine $U/\{ud\}$ in the new context, we need assign each object o_j to the right tolerance class X_i representing the decision value ud_i. We should check for each object o_j the distance

Table I
UNCERTAIN DECISION TABLE

U	Hair	Eyes	Height	ud
o_1	Dark	Brown	Short	$m_1(\{ud_2\}) = 0.5$ $m_1(\Theta) = 0.5$
o_2	Blond	Blue	Middle	$m_2(\{ud_2\}) = 1$
o_3	Blond	Brown	Short	$m_3(\{ud_1\}) = 0.7$ $m_3(\Theta) = 0.3$
o_4	Blond	Brown	Tall	$m_4(\{ud_1\}) = 0.9$ $m_4(\Theta) = 0.1$
o_5	Dark	Brown	Short	$m_5(\{ud_2\}) = 1$
o_6	Blond	Blue	Middle	$m_6(\{ud_2\}) = 0.9$ $m_6(\Theta) = 0.1$
o_7	Dark	Brown	Tall	$m_7(\{ud_1\}) = 1$
o_8	Dark	Brown	Middle	$m_8(\{ud_1\}) = 0.975$ $m_8(\Theta) = 0.025$

between its bba m_j and the decision value ud_i. The idea is to use a distance measure between the two bba's m_j and a certain bba m (such that $m(\{ud_i\}) = 1$). Many distance measures between two bba's were developed [2], [3], [4], [17]. In our case, we choose the distance measure proposed in [4] which is directly defined on bba's and satisfies more properties such as non-negativity, non-degeneracy and symmetry.

For every ud_i, we define a tolerance class as follows:

$$X_i = \{o_j | dist(m, m_j) < 1 - threshold\} \quad (10)$$

such that $m(\{ud_i\}) = 1$. Besides, we define a tolerance relation as follows:

$$IND_{\{ud\}} = U/\{ud\} = \{X_i | ud_i \in \Theta\} \quad (11)$$

It should be noted here that we have replaced the term equivalence class from the certain decision attribute case by tolerance class for the uncertain decision attribute, because the resulting classes may overlap.

Example: *Let us continue with the same example to compute the equivalence classes based on condition attributes in the same manner as in the certain case: U/C=* $\{\{o_1\}, \{o_2, o_6\}, \{o_3\}, \{o_4, o_5\}, \{o_7\}, \{o_8\}\}$ *and to compute the tolerance classes based on the uncertain decision attribute* $U/\{ud\}$*. If we take threshold equal to* 0.1*, we obtain the following results:*

For the uncertain decision value ud_1*, we obtain* X_1 =$\{o_1, o_3, o_4, o_7, o_8\}$*. For the uncertain decision value* ud_2*, we obtain* X_2 =$\{o_1, o_2, o_3, o_5, o_6\}$*.* $U/\{ud\}$=$\{\{o_1, o_3, o_4, o_7, o_8\}, \{o_1, o_2, o_3, o_5, o_6\}\}$.

3) Set approximation: Like the tolerance relation $U/\{ud\}$, the set approximation concept should be redefined in the new uncertain context. Hence, to compute the new $lower$ and $upper$ approximations for our UDT, we should follow two steps:

1) We combine the bba's for each equivalence class from U/C using the mean operator [7] which is more suitable than the rule of combination in eqn. (9) which is proposed especially to combine different beliefs on decision for one object and not different bba's for different objects as follows:

$$\bar{\text{m}}_{[o_j]_C}(E) = \frac{1}{|[o_j]_C|} \sum_{o_i \in [o_j]_C} m_i(E), \ for\, all\; E \subseteq \Theta \quad (12)$$

Let us remind that $[o_j]_C$ is the equivalence class containing the object o_j.

2) We compute the new $lower$ and $upper$ approximations for each tolerance class X_i from $U/\{ud\}$ based on uncertain decision attribute ud_i as follows:
$\underline{C}(X_i) =$

$$\{o_j | [o_j]_C \subseteq X_i \; and \; dist(m, \bar{\text{m}}_{[o_j]_C}) \leq threshold\} \quad (13)$$

$$\bar{C}(X_i) = \{o_j | [o_j]_C \cap X_i \neq \emptyset\} \quad (14)$$

We find in the new $lower$ approximation all equivalence classes from U/C included in X_i where the distance between the combined bba $\bar{\text{m}}_{[o_j]_C}$ and the certain bba m (such that $m\{ud_i\}) = 1$) is less than a $threshold$. However, the $upper$ approximation is computed in the same manner as in the certain case. The boundary region of X_i stills have also the same definition:

$$BN_C(X_i) = \bar{C}(X_i) - \underline{C}(X_i) \quad (15)$$

Note that in the case of uncertainty the $threshold\ value$ gives more flexibility to the tolerance relation and the set approximations.

Example: *We continue with the same example to compute the new lower and upper approximations. After the first step, we obtain the combined bba for each equivalence class from* U/C *using mean operator. Next, we compute the lower and upper approximations for each tolerance class* $U/\{ud\}$*. We will use* $threshold = 0.1$*.*

For the uncertain decision value ud_1, let X_1 =$\{o_1, o_3, o_4, o_7, o_8\}$. We obtain: $\underline{C}(X_1)$=$\{o_4, o_7, o_8\}$, $\bar{C}(X_1)$=$\{o_1, o_3, o_4, o_5, o_7, o_8\}$ and $BN_C(X_1)$=$\{o_1, o_3, o_5\}$.

For uncertain decision value ud_2, let X_2 =$\{o_1, o_2, o_3, o_5, o_6\}$. We obtain: $\underline{C}(X_2)$=$\{o_2, o_6\}$, $\bar{C}(X_2)$=$\{o_1, o_2, o_3, o_5, o_6\}$ and $BN_C(X_2)$=$\{o_1, o_3, o_5\}$.

4) Positive region and dependency of attributes: With this new $lower$ approximation, we can define the new positive region denoted $UPos_C(A, \{ud\})$:

$$UPos_C(A, \{ud\}) = \bigcup_{X_i \in U/\{ud\}} \underline{C}(X_i) \quad (16)$$

The attribute ud depends on the set of attributes C in a degree k, if

$$k = \gamma_C(A, \{ud\}) = \frac{|UPos_C(A, \{ud\})|}{|U|} \quad (17)$$

Example: *Let us continue with the same example, to compute the positive region and dependency of attributes of A.*

$$UPos_C(A, \{ud\}) = \{o_2, o_4, o_6, o_7, o_8\}$$

$$\gamma_C(A, \{ud\}) = \tfrac{5}{8}$$

B. Construction procedure of BRSC based on dynamic core

Our technique of classification within the belief function framework denoted the Belief Rough Set Classifier was originally proposed in [16]. This technique is based on the basic concepts of rough sets and has the advantages of time complexity of learning, accuracy of classification and size of discovered rules. To overcome the problem of noise existing in real databases, the construction procedure of the BRSC based on dynamic core is described by means of the following steps:

Step 1. Eliminate the superfluous condition attributes: We remove the superfluous condition attributes that are not in reducts. This leaves us with a minimal set of attributes that preserve the ability to perform the same classification as the original set of attributes. We need to redefine the concepts of reduct and core in this new situation. Using the new formalism of positive region, the new definition of reduct is a minimal set of attributes $B \subseteq C$ such that:

$$UPos_B(A, \{ud\}) = UPos_C(A, \{ud\}) \quad (18)$$

The core is the most important subset of attributes, is intersection of all reducts.

$$UCore_C(A, \{ud\}) = \bigcap URed_C(A, \{ud\}) \quad (19)$$

Where $URed_C(A, \{ud\})$ is the set of all reducts of A relative to ud. Nevertheless, computing reducts from uncertain and noisy data leads to results which are unstable and sensitive to the sample data. Therefore it is important to search the most stable reduct denoted dynamic reduct [15] or computing reduct containing dynamic core defined as follows:

If $A = (U, C \cup \{ud\})$ is an uncertain decision table, then any system $A' = (U', C \cup \{ud\})$ such that $U' \subseteq U$ is called a subtable of A. Let F be a family of subtables of A.

$$DCore_C(A, F) =$$

$$UCore_C(A, \{ud\}) \cap \bigcap_{A' \in F} UCore_C(A', \{ud\}) \quad (20)$$

Any element of $DCore_C(A, F)$ is called an F-dynamic core of A. From the definition of dynamic core, it follows that a relative core of A is dynamic if it is also a core of all subtables from a given family F.

Example: *Let us continue with the same example in Table I to compute the possible reducts using the new formalism of positive region.*

$UPos_{\{Hair, Height\}}(A, \{ud\}) = UPos_C(A, \{ud\})$
$UPos_{\{Eyes, Height\}}(A, \{ud\}) = UPos_C(A, \{ud\})$

There are two reducts: $\{Hair, Height\}$ and $\{Eyes, Height\}$. The attribute $Height$ is the only core. However, we should check that $Height$ is also the core of any chosen sub-decision table from A. If it is the case, we have two possible solutions to simplify our uncertain decision table. If we choose the second reduct $\{Eyes, Height\}$, the UDT is simplified as shown in Table II.

Table II
THE SECOND REDUCT

U	Eyes	Height	ud
o_1	Brown	Short	$m_1(\{ud_2\}) = 0.5 \quad m_1(\Theta) = 0.5$
o_2	Blue	Middle	$m_2(\{ud_2\}) = 1$
o_3	Brown	Short	$m_3(\{ud_1\}) = 0.7 \quad m_3(\Theta) = 0.3$
o_4	Brown	Tall	$m_4(\{ud_1\}) = 0.9 \;\; m_4(\{ud_2\}) = 0.1$
o_5	Brown	Short	$m_5(\{ud_2\}) = 1$
o_6	Blue	Middle	$m_6(\{ud_2\}) = 0.9 \;\; m_6(\Theta) = 0.1$
o_7	Brown	Tall	$m_7(\{ud_1\}) = 1$
o_8	Brown	Middle	$m_8(\{ud_1\}) = 0.975 \; m_8(\Theta) = 0.025$

Step 2. Eliminate the redundant objects: After removing the superfluous condition attributes, we will find redundant objects (having the same condition attribute values). They may not have the same bba on decision attribute. So, we use their combined bba's based on the mean operator as a rule of combination as follows:

$$\bar{m}_{[o_j]_B}(E) = \frac{1}{|[o_j]_B|} \sum_{o_i \in [o_j]_B} m_i(E), \;\; for\, all\, E \subseteq \Theta \quad (21)$$

Where B is the relative reduct of C with respect to ud.

Example: *After removing the superfluous condition attributes and the redundant objects for the uncertain decision table, we obtain Table III. Let us remind that we have used the notation $m_{1,3,5}$ to mean $\bar{m}_{[o_1]_B}$ with B is equal to {Eyes, Height}.*

Table III
COMBINED BBA'S OF REDUNDANT OBJECTS

U	ud
o_1, o_3, o_5	$m_{1,3,5}(\{ud_1\}) = 0.24 \; m_{1,3,5}(\{ud_2\}) = 0.5$ $m_{1,3,5}(\Theta) = 0.26$
o_2, o_6	$m_{2,6}(\{ud_2\}) = 0.975 \; m_{2,6}(\Theta) = 0.025$
o_4, o_7	$m_{4,7}(\{ud_1\}) = 0.975 \;\; m_{4,7}(\{ud_2\}) = 0.025$
o_8	$m_8(\{ud_1\}) = 0.975 \; m_8(\Theta) = 0.025$

Step 3. Eliminate the superfluous condition attribute values: To further simplify the uncertain decision table,

we can eliminate some attribute values. We also need to redefine the concept of value reduct for each decision rule of the form: **If** $C(o_j)$ **then** m_j as follows:

For all $B \subset C$, Let $X = \{o_k | B(o_j) = B(o_k)\, and\, j \neq k\}$
If $X = \emptyset$ **then** $B(o_j)$ is a value reduct of o_j.
Else If $Max(dist(m_j, m_{[o_k]_C})) \leq$ threshold **then** $B(o_j)$ is a value reduct of o_j.

Example: *If we compute the value reduct for all belief decision rules, we obtain Table IV.*

Table IV
SIMPLIFIED UNCERTAIN DECISION TABLE

U	Eyes	Height	ud
o_1, o_3, o_5	Brown	Short	$m_{1,3,5}$
o_2, o_6	Blue		$m_{2,6}$
o_4, o_7		Tall	$m_{4,7}$
o_8	Brown	Middle	m_8

1) Generation of belief decision rules: The decision rules induced from our uncertain decision table are denoted belief decision rules where the decision is represented by a bba. Hence, these rules are simplified by removing superfluous condition attributes and condition attributes values. With simplification, we can improve the time and the performance of classification of unseen objects.

Example: *Table IV gives one solution of rule classification described as follows:*
If Eyes=Brown and Height=Short then $m_{1,3,5}(\{ud_1\}) = 0.24$ $m_{1,3,5}(\{ud_2\}) = 0.5$ $m_{1,3,5}(\Theta) = 0.26$
If Eyes=Blue then $m_{2,6}(\{ud_2\}) = 0.975$ $m_{2,6}(\Theta) = 0.025$
If Height=Tall then $m_{4,7}(\{ud_1\}) = 0.975$ $m_{4,7}(\{ud_2\}) = 0.025$
If Eyes=Brown and Height=Middle then $m_8(\{ud_1\}) = 0.975$ $m_8(\Theta) = 0.025$

V. STUDY DATA AND EXPERIMENTAL RESULTS

In this Section, we will report the experimental results by applying our technique of classification BRSC on web usage mining dataset.

A. Data Description

The study data were obtained from web access logs of the introductory computing science course at Saint Mary's University. The course is "Introduction to Computing Science and Programming" offered in the first term of the first year. The initial number of students in the course was 180. The number reduced over the course of the semester to 140 students. The students in the course come from a wide variety of backgrounds, such as computing science major hopefuls, students taking the course as a required science course, and students taking the course as a science or general elective. As is common in a first year course, students' attitudes towards the course also vary a great deal. Lingras and West [5] showed that visits from students attending the first course could fall into one of the following three categories (decision values):

- Studious: These visitors download the current set of notes. Since they download a limited/current set of notes, they probably study class-notes on a regular basis.
- Crammers: These visitors download a large set of notes. This indicates that they have stayed away from the class-notes for a long period of time. They are planning for pretest cramming.
- Workers: These visitors are mostly working on class or lab assignments or accessing the discussion board.

The web logs were preprocessed to create an appropriate representation of each user, corresponding to a visit. The abstract representation of a web user is a critical step that requires a good knowledge of the application domain. Based on some observations, it was decided to use the following attributes for representing each visitor [5]:

- On campus/Off campus access.
- Day time/Night time access: 8 a.m. to 8 p.m. were considered to be the daytime.
- Access during lab/class days or non-lab/class days: All the labs and classes were held on Tuesdays and Thursdays. The visitors on these days are more likely to be workers.
- Number of hits.
- Number of class-notes downloads.

The first three attributes had binary values of 0 or 1. The last two variables represent the number of hits and number of class-notes, and were integer values. Total visits were 23754. The visits where no class-notes were downloaded were eliminated, since these visits correspond to either casual visitors or workers. Elimination of outliers and visits from the search engines further reduced the sizes of the data set to 7965. Instead of representing an object as belonging to a cluster ud_i (decision value), we also associated a degree of belief in the object belonging to the cluster ud_i. Let $\Theta = \{ud_i | 1 \leq i \leq k\}$ be the set of all the clusters (the possible decision values). If K-means algorithm identified the object o_j as belonging to cluster ud_i, then the bba m_j for the object was calculated as follows:

$$m_j(\{ud_i\}) = \frac{sim(o_j, ud_i)}{\sum_{j=1}^{k} sim(o_j, ud_i)}, m_j(\Theta) = 1 - m_j(\{ud_i\}) \tag{22}$$

Where $sim(o_j, ud_i) = 1/d(o_j, ud_i)$ represents the similarity between object o_j and ud_i and is calculate as the inverse of the distance d between them. The latter is a standard Euclidean distance.

B. Evaluation criteria

The relevant criteria used to judge the performance of our classifier to create rules from the web usage mining database are as follows:

1) *The accuracy* represents the percent of correct classification (PCC) of the objects belonging to testing set.
2) *The size* represents the number of the belief decision rules generated from the classifier.
3) *The time requirement* represents the number of seconds needed for the construction procedure.

To more evaluate our tests, we compare the results with those obtained from similar classifier denoted Belief Decision Tree [1]. The latter can generate classification decision rules from our web usage database.

Table V
THE EXPERIMENTAL RESULTS

Approaches	PCC (%) certain case	PCC (%) uncertain case	Size	Time requirement
BRSC	85.24	89.63	37	113
BDT	84.07	85.12	41	108

Table V summarizes the results relative to the three chosen evaluation criteria obtained by applying the BRSC to learn decision rules from the web usage mining. We find that PCC is equal to 85.24 % based on the certain cluster assigns to each object. However, the PCC becomes 89.63% based on the degree of beliefs associated to each cluster. The size of generated model is 37. The time requirement is corresponding to 113 seconds. The size and the time requiement are almost the same for the ceratin and uncertain cases. So, we conclude from the positive obtained results that it is interesting to apply our BRSC to web mining databases characterized by uncertain decision values by the means of belief functions more than the BDT.

VI. CONCLUSION

In this paper, a new approach of classification system based on rough sets named BRSC have been applied to generate a classification model from uncertain data consisting of web usage. The uncertainty appears only in decision attributes and is handled by the TBM, one interpretation of the belief function theory. The feature selection step used to construct the BRSC is based on the calculation of dynamic core to extract more relevant and stable features for the classification process. In experimentations, three evaluation criteria have been chosen to judge the performance of the BRSC applied to the web usage mining dataset. We find interesting results that may encourage users or experts in the web domain to use our BRSC to handle uncertainty in the decision attribute. As a future work, we hope to handle the problem of uncertainty in condition attribute values.

REFERENCES

[1] Elouedi, Z., Mellouli, K., Smets, P.: Belief decision trees: Theoretical foundations. International Journal of Approximate Reasoning, Vol 28(2-3), 91124 (2001)

[2] Elouedi, Z., Mellouli, K. and Smets, P.: Assessing sensor reliability for multisensor data fusion within the transferable belief model. IEEE Trans Syst Man cyben, Vol 34(1), (2004) 782-787.

[3] Fixen, D. and Mahler, R.P.S.: The modified Dempster-Shafer approach to classification, IEEE Trans Syst Man Cybern. Vol 27(1), (1997) 96-104.

[4] Jousseleme, A.L., Grenier, D. and Bosse, E.: A new distance between two bodies of evidence. Information Fusion. Vol 2(2), (2001) 91-101.

[5] Lingras, P. and West, C.: Interval Set Clustering of Web Users with Rough K-means, Journal of Intelligent Information Systems, Vol 23(1), 2004(5-16).

[6] Maheswari, V. U., Siromoney, A. and Mehata, K. M.: The variable precision rough set model for web usage mining, Proc. 1st Asia-Pacific Conf.Web Intell. WI-2001, Maebashi, Japan, Oct. 2001.

[7] Murphy, C.K.: Combining belief functions when evidence conflits, Decision Support Systems 29, (2000) 1-9.

[8] Pawlak, Z.: Rough Sets. International journal of computer and information sciences, Vol 11(5), (1982) 341-356.

[9] Pawlak, Z.: Rough Sets: Theoretical Aspects of Reasoning About Data. Dordrecht: Kluwer Academic Publishing, 1991.

[10] Shafer, G.: A mathematical theory of evidence. Princeton University Press. Princeton, NJ, 1976.

[11] Smets, P. and Kennes, R.: The transferable belief model. Artificial Intelligence, Vol 66 (2), (1994) 191-234.

[12] Smets, P.: The transferable belief model for quantified belief representation. In D.M. Gabbay & P. Smets (Eds), Handbook of defeasible reasoning and uncertainty management systems, Vol 1, Doordrecht, The Netherlands: Kluwer, (1998) 267-301.

[13] Smets, P.: Application of the transferable belief model to diagnostic problems. International journal of intelligent systems, Vol 13 (2-3), (1998) 127-157.

[14] Trabelsi, S. and Elouedi, Z.: Learning decision rules from uncertain data using rough sets, The 8th International FLINS Conference on Computational Intelligence in Decision and Control, Madrid, Spain, September 21-24, World scientific, (2008) 114-119.

[15] Trabelsi, S., Elouedi, Z. and Lingras, P.: Dynamic reduct from partially uncertain data using rough sets, RSFDGrC 2009, LNAI 5908, (2009) 160-167.

[16] Trabelsi, S., Elouedi, Z. and Lingras, P.: Belief rough set classifier, Canadian AI 2009, LNAI 5549, (2009) 257-261.

[17] Zouhal, L.M and Denoeux, T.: An evidence-theory k-NN rule with parameter optimization, IEEE Trans.Syst.Man Cybern. Vol 28(2), (1998) 263-271.