1. Introduction

The World-Wide Web\textsuperscript{1–3} forms a big source of information among various domains. The scalability of data makes the database volume wider and wider. However, the era of the mobile internet still increases the data volume. The internet society\textsuperscript{4} quotes that by May 2015 there were 3 billion internet users and it’s likely to increase 71% in 2019. The data is distributed globally and most of the internet users really on this data. Hence it increases the complexity of searching, managing and analyzing of those data. Moreover, there is not much semantics associated with the data in the WWW, makes the process more human interaction than computer interaction. The Semantic Web\textsuperscript{5,6} is a vision of the next generation World-Wide Web in which data from multiple sources described with rich semantics are integrated to enable processing by humans as well as software agents. One of the goals of Semantic Web research is to incorporate most of the knowledge of a domain in an ontology that can be shared by many applications. Ontologies\textsuperscript{7,8} provide hierarchal taxonomies of information of a particular domain with concepts based classes, attributes, and the relationships between concepts.

Today most of the web pages are designed for humans to read and get information, rather than the intellectual processing by the computer\textsuperscript{9}. The web pages are designed according to the defined structure which is understandable only to the humans. For example, if a student needs to know about MBBS course, the current web will give only the details of the MBBS course. Instead, the web provides the details of the course, where it is offered, what is the duration, payment, etc. then it becomes the beneficial process. This could be done through the semantic web. The semantic web is just an extension of the current web, which processes the annotated semantics along with the data being stored on the web. The Semantic Web will bring structure to the meaningful content of Web pages, creating an environment where software agents roaming from page to page can readily carry out sophisticated tasks for users.

On the extension of above definition, the semantic web had been revisited and put into existence\textsuperscript{10}. The standards organizations like the Internet Engineering Task Force and the World Wide Web Consortium (W3C), has directed major efforts at specifying, developing, and deploying languages for sharing meaning. These languages provide a foundation for semantic interoperability. In
1997, the W3C defined the first Resource Description Framework (RDF) specification. RDF provided a simple but powerful triple-based representation language for Universal Resource Identifiers (URIs). It became a W3C recommendation by 1999. The vision for RDF was to provide a less essential knowledge representation for the Web.

2. Rough Set

Rough set theory\(^1\) was proposed by as a new mathematical tool for the imprecise data. The indiscernibility among the objects associated with the same category leads the concept of approximation. Figure 1 shown as the basis of the rough set which formulates the definition of two sets as lower and upper approximation.

![Rough set approximations](image)

Figure 1. Rough set approximations.

Let \( U \) be the universe and \( R \) be the equivalence relation defined on \( U \), if \( A \) is the target set then \( A \) can be characterized by \( RA \) (lower approximation) and \( RA^c \) (upper approximation).

\[
RA = \bigcup \{ Y \in U / R : Y \subseteq A \}
\]

\[
RA^c = \bigcup \{ Y \in U / R : Y \cap A \neq \phi \}
\]

Rough set theory is an effective and powerful methodology and has been widely used in various domains for knowledge discovery. Many people have discovered knowledge by combining various mining and intelligent techniques with the rough set. For example in the medical domain knowledge has been discovered by using the rule-based induction algorithms\(^2\) and build a knowledge-based system to detect the early ovarian cancer by analysing the big data\(^3\).

3. Ontology on Intelligent Techniques

The ontology system in the semantic web can be constructed based on Formal Concept Analysis (FCA) and Rough set\(^4\). Today web-surfing becomes a day-to-day activity where people search for information. The ontology provides the explicit specification of the domain knowledge in the semantic web. This makes the clear description of the knowledge and makes the surfing at ease with the means of a semantic network. The essential and critical part in the construction of the ontology conceptual model is to identify the concepts and the relationship between those concepts. Even though many methodologies have been developed to build the ontology, the construction is still difficult as the object orientation model is always an abstraction to the ontology.

To construct ontology requires more costs, works, efforts, and domain knowledge. The users were caught up by the annotation used to extricate the interdependencies in the semantic web.

As a new model of set-theoretical for the concept and conceptual hierarchies the Formal Concept Analysis (FCA) has been proposed by\(^5\). This model not only forms the basis for data analysis but for also to represent the conceptual knowledge and information management. An Intelligent Framework for medical diagnosis was also designed by\(^6\). Domain Knowledge or true belief can be expressed in terms of concepts which represent a set of objects or its instance which share the similar characteristics Formal Concept Analysis allow a mathematical description of a domain with the formal context triplet \( K = (X, Y, R) \), where \( G \) is a non-empty set of objects, \( M \) is a non-empty set of attributes and \( R \) a binary relation between the objects and the attributes i.e \( R \subseteq X \times Y \). A concept is determined by its extent which is a set of objects and its intent which is a set of attributes shared by the objects. For every \( A \subseteq X \) and for every \( B \subseteq Y \), the formal context can be defined by the formulas

\[
\forall A, \subseteq A = \{ y \in B | \forall x \in A, x Ry \}
\]

\[
\forall B, \subseteq B = \{ y \in A | \forall y \in B, x Ry \}
\]

Formal Context can be expressed in terms of the table. If \( \{x_1, x_2, x_3, \ldots, x_m\} \) are the set of objects and \( \{y_1, y_2, y_3, \ldots, y_n\} \) are the set of attributes with \( m \) rows and \( n \) columns,
then the relation R defined by \((x_i, y_j) \in R\) only if the Table 1 shown as \(i\) th row and \(j\) th column contains the value \(x\).

For example:

<table>
<thead>
<tr>
<th>Table 1. Object-attribute matrix</th>
</tr>
</thead>
<tbody>
<tr>
<td>R</td>
</tr>
<tr>
<td>X1</td>
</tr>
<tr>
<td>X2</td>
</tr>
<tr>
<td>X3</td>
</tr>
<tr>
<td>X4</td>
</tr>
</tbody>
</table>

Concept lattice is a form of an ordered taxonomical graph in which each node represents a subset of objects (intent) with their common attributes (extent) and the intent and the extent of a formal concept uniquely determine each other. For this case, attributes are called many-valued attributes. Since Formal Concept analysis presents multivalued attributes this can be best used for the semantic web to express the domain knowledge. Moreover, the ontology is constructed based on classes; FCA can be used to construct the ontology in an effective way.

Formal Concept Analysis can be done automatically by means of computer programs. This makes the task of data analysing and knowledge representation in ease. Hence the FCA can be used to construct the ontology easier and effective.

Finally, the author introduces the rough concept analysis by including rough set to the Formal Concept Analysis. The Rough Set theory introduces by addresses the vagueness, imprecision, and uncertainty in data analysis. In the semantic web, ontology is constructed and reused. The ontology may be developed, aligned, mapped, and merged during the process. During such time the Rough Set can be used to meet the boundary conditions. The author had developed an architecture where data is extracted from the log file with which the formal contexts are developed. Later a lattice model is constructed with the help of rough set which is stored as the domain ontology.

The information retrieval can be effectively done based on rough ontology\(^1\). In the traditional keyword-based search method the word sense disambiguation leads to lack of accuracy. The limitations of the traditional keyword-based search are overwhelmed with the advent of the semantic web. With the semantic annotation, ontology plays a vital role in the field of information retrieval. As the ontology provides the formal explicit specification of conceptualization, it makes the domain knowledge machine understandable. Though the ontology affords the semantic annotation, unfortunately, it doesn’t address the imprecise, uncertain and vague concept of real life data. The dominant ontology theory is further enriched with other powerful Rough set theory to process the imprecise information. In the proposed method the information retrieval is enhanced with the rough ontology based model. The proposed model is precisely defined by the following procedure.

- Keyword Query (Q) is posted by the user.
- Using the traditional keyword-based method, Initial document set (Dq) is obtained.
- The upgradation of actual documents to the machine understandable is done through the semantic annotation by constructing the specified domain ontology.
- To the annotated document association search is made and a new individual set (Iq) and property set (Pq) are obtained.
- Having Pq as the equivalence relation an approximation space is constructed.
- Taking the individual Dq and Iq the similarity is computed based on the defined approximation space.
- According to the similarity ranking is provided and the final search document set is sent to the user.

The authors also proved that the performance of their proposed method was upright with the traditional keyword-based method and ontology-based query retrieval method using the weighted harmonic mean of precision and recall.

Case Representation is a way to represent the knowledge. In\(^1\)\(^8\) presented their views on to present the knowledge by means of ontology-based case representation using rough set. Case-based Reasoning was first formulated by\(^1\)\(^9\). In our day to day life, we can see people solving their problems using their previous knowledge and experience. The same methodology is followed in case representation where we derive the new solutions based on the old experiences. Normally this type methodology can be adapted to situations like problem solving and interpretation where we recall the old experience and adapt them to the new solution or interpret to new solutions.

The Traditional way to represent the case in case based reasoning are mainly classified into three categories as feature vector representation, structured representation...
and textual representation. Later this is extended by hierarchical and general cases\textsuperscript{20}. In addition each case in the case library in case based reasoning using the domain ontology\textsuperscript{21}. Since there is lack of flexibility in process of retrieval Qian Chen et al. uses rough set to define the clustering.

The most fundamental concept of rough set is the approximation. A view of rough set concept approximation is given by \textsuperscript{22}. Some of the basic approaches determining the rough approximation to a concept are attribute-based approximations, case-based approximations, and rule-based approximations. An improved method of classifier synthesis based on rough set and layered learning approach has also been studied\textsuperscript{23}. The authors compared their current work with layered learning algorithm with the classical RS algorithm for concept approximation. Using the RS algorithm the decision attribute of target attribute is connected directly with the input attributes.

The main problem of this method is the construction of the decision table and to the lack of decision attribute for intermediate concepts. The work is improved by using the supervised learning concept where the decision attribute of the target concept is introduced as class attributes. Moreover, he used an ontology to represent the domain knowledge where the concepts are represented in a hierarchical approach. In the proposed method the training set is represented by a decision table and the set of decision attributes corresponds to both intermediate and target attributes. The improved layered learning algorithm takes the decision system and the concept hierarchy as input and gives the hypothetical attributes of all concepts in the hierarchy as output. The algorithm at each level takes the concepts and check with the input attribute set if present it adds to the final set else it forms the concept approximation through the rough set theory.

4. Conclusion

This paper probes how the rough ontology can be applied to the indiscernibility of objects. The vague and incomplete knowledge in the semantic web can be managed by the rough set. Although managing ontology in the approximation space is difficult it can be overcome by using different intelligent techniques as it has been discussed in this paper.

5. References