

A Real Time Application of Soft Set in Parameterization Reduction for Decision Making Problem

Janmejy Pant¹, Amit Juyal², Himanshu Pant³, Akhilesh Dwivedi⁴

^{1,3,4}Department of Computer Science and Applications, Graphic Era Hill University, Bhimtal Campus, India

²Department of Computer Applications, Graphic Era Hill University, Dehradun Campus, India

Article Info

Article history:

Received Jul 2, 2016

Revised Nov 28, 2016

Accepted Dec 11, 2016

Keyword:

Approximation

Data uncertainty

Information system

Parameterization

Reduct

Soft set predicate

Softset theory

ABSTRACT

In each and every field of science and technology Information science plays an important role. Sometimes information science is facing different types of problems to handle the data and information. Data Uncertainty is one of the challenging difficulties to handle. In past, there are several theories like fuzzy set, Rough set, Probability etc. to dealing with uncertainty. Soft set theory is the youngest theory to deal with uncertainty. In this paper we discussed how to find reducts. This paper focuses on how we can transform a sample data set to binary valued information system. We are also going to reduce the dimension of data set by using the binary valued information that results a better decision.

Copyright © 2017 Institute of Advanced Engineering and Science.
All rights reserved.

Corresponding Author:

Janmejy Pant,

Departement of Computer Science and Applications,

Graphic Era Hill University,

Sattal Road, PO: Bhowali, Bhimtal, Nainital (Uttarakhand), India.

Email: geujay2010@gmail.com

1. INTRODUCTION

There are many fields like economics, engineering, environment, social science and medical science which involve data of uncertainties. The traditional mathematical tools are not able to handle such type of uncertainties. The researchers used the existing traditional tools such as probability theory, Fuzzy set theory, [9-11] theory of mathematics, rough set theory etc. But these theories are not much successful to deal with uncertainty because they all have their own difficulties. One reason for these difficulties may be due to the inadequacy of the parameterization tools [1]. Molodtsov [2] developed the concept of soft set theory as an effective mathematical tool for dealing with uncertainties and non-crisp data. Soft set theory is free from the difficulties that have troubled the usual theoretical approaches. He pointed out several directions for the applications of soft set. Soft Set is called (binary, basic, elementary) neighborhood systems [3]. The soft set [10], [11] is a mapping from parameter to the crisp subset of universe. The structure of soft set can classify the objects into two classes such as yes/1/true or no/0/false [4]. This statement shows that a standard soft set deals with Boolean valued systems. Method to describe the step of research and used in the chapter "Results and Discussion" to support the analysis of the results [2]. If the manuscript was written really have high originality, which proposed a new method or algorithm, the additional chapter after the "Introduction" chapter and before the "Research Method" chapter can be added to explain briefly the theory and/or the proposed method/algorithm [4].

2. SOFT SET FUNDAMENTALS

In this section soft set theory fundamentals are explained.

2.1. Information System in Soft Set

In relational database relations are known as tables and a relation is the combination of rows and columns. In the same way information systems are similar to the relations on relational database. An information systems of a soft set is a quadruple i.e. 4-tuples [1], [4], [5]

$$S = (U, A, V, f) [1]$$

where,

U-Non empty finite set of objects

Non empty finite set of attributes

V- Value set of attributes

f- Information function

An information System is also known as knowledge representation system or an attribute valued system.

2.2. Introduction to Softset Theory

Soft set is a simple mathematical tool that deals with a collection of approximate behavior or description of objects. It is a parameterized tool to deal with approximate value and uncertainty. Each approximate description can divide in two parts. First one is a predicate part and second one is an approximate value set [1]. In classical mathematics, a mathematical model of an object is constructed and defines the notion of an exact solution of this model. A classical model is generally used to find the exact solutions i.e. this type of model is used for precise value. In mathematical model it is a tough task to obtain the exact solution for the imprecise values. So, it is more reliable and easy to define an approximate solution first and then the solution is calculated. The soft set theory has given an opposite approach to solve these types of problems. In the initial description of an object it has an approximate value and it does not required to introduce the notion of exact solution. These types of conditions on approximate behavior are not present in soft set theory. This property makes soft set theory reliable, simple and comfortable to apply in any real world applications. With the help of words, sentences, numbers, functions the parameterization of soft set can be expressed.

In the whole paper U indicates to the universe, E is defined by a set of parameters and P (U) is the power set of U.

Definition 1. ([1], [4]). A pair (F, E) is called a soft set over U, where F is a mapping given by

$$F: E \rightarrow P(U)$$

In other words U is universe in soft set and it is a parameterized family of subsets of universe U. for $e \in E$, $F(e)$ may be considered as the set of e elements of the soft set (F, E) as the set of e approximate elements of soft set [1]. Ultimately, soft set is a non-crisp set

Definition 2. (See [1]). Let R be a family of equivalence relations and let $A \in R$. It is say that A is dispensable in R if $IND(R) = IND(R - \{A\})$ [1-2], [4]. Otherwise A is indispensable in R. The family R is Independent if each $A \in R$ is indispensable in R. otherwise R is dependent. Q subset of P is a reduction of P if Q is independent and $IND(Q) = IND(P)$, that is to say Q is the minimal subset of P that keeps the classification ability [7]. The set of all indispensable relations in P is called the core of P, and is denoted as CORE (P). Clearly, $CORE(P) = \cap RED(P)$, where RED (P) is the family of all reductions of P [2], [7], [10].

Definition 3. (See [1], [8]). Let (F, E) be a soft set over the universe U and $u \in U$. A parameter co-occurrence set of an object u can be defined as [8]:

$$Coo(u) = \{e \in E: f(u, e) = 1\}$$

Obviously,

$$Coo(u) = \{e \in E: F(e) = 1\}$$

Definition 4. (See [1]). Let (F, E) be a soft set over the universe U and $u \in U$.

Support of an object u is defined by:

$$\text{Supp}(u) = \text{card}(\{e \in E: f(u, e) = 1\}) [1]$$

3. EXAMPLE OF REDUCTION IN SOFT SET

In this example there is soft set (F, E) which describes the “nature of Bikes to purchase” that a customer is supposed to make a necessary action as a decision to buy. In the showroom there are five types of bikes are ready to sale. The nature of different types of bikes is Small bike, Medium bike, Expensive bike, Second-hand bike, and imported bike.

Suppose we have six bikes which are under consideration,

$$U = \{b1, b2, b3, b4, b5, b6\}$$

where, $b1, b2, b3, b4, b5, b6$ are bikes under consideration and E is a set of decision parameters,

$$E = \{e1, e2, e3, e4, e5\} [1]$$

where,

For the parameter “small bike”, we have $e1$

For the parameter “medium bike”, we have $e2$

For the parameter “expensive bike”, we have $e3$

For the parameter “Second-hand bike”, we have $e4$

For the parameter “imported bike”, we have $e5$

Consider the mapping $F: E \rightarrow P(U)$ given by “bikes $(.)$ ”, Here $(.)$ is to be filled in by one of parameters $e \in E$ [1]

Suppose that

$$F(e1) = \{b2, b3, b4, b5\},$$

$$F(e2) = \{b1, b6\},$$

$$F(e3) = \{b1, b2, b6\},$$

$$F(e4) = \{b1, b2, b3, b4, b5, b6\},$$

$$F(e5) = \{b1, b2, b3, b4, b5, b6\}.$$

Therefore, $F(e1)$ means “the size of bikes are small”, its functional value is the set is $F(e1) = \{b2, b3, b4, b5\}$. It is the soft set (F, E) as a collection of approximations as:

$$(F, E) = \left(\begin{array}{l} \text{small bike} = \{b2, b3, b4, b5\} \\ \text{Medium bike} = \{b1, b6\} \\ \text{Expensive bike} = \{b1, b2, b6\} \\ \text{Second hand bike} = \{b1, b2, b3, b4, b5, b6\} \\ \text{Imported bike} = \{b1, b2, b3, b4, b5, b6\} \end{array} \right)$$

The soft set (F, P) is can be expressed as a binary table, as shown Table 1, to solve this problem [1]. For this $h_{ij} = 1$ if $h_i \in F(e_j)$ then $h_{ij} = 1$, otherwise $h_{ij} = 0$, where h_{ij} are the entries in Table 1.

So a soft set can now be known as a knowledge representation system. In this system we generally use set of parameters instead of set of attributes. Every approximation can divided into two parts, first a predicate part e and second one is approximate value set p [1]. for example, for the approximation small bike = $\{b2, b3, b4, b5\}$, have the predicate name of bikes with small size and the value set is $\{b2, b3, b4, b5\}$ [1].

Table 1. Representation of soft set with the above example in table

U	$e1$	$e2$	$e3$	$e4$	$e5$
b1	0	1	1	1	1
b2	1	0	1	1	1
b3	1	0	0	1	1
b4	1	0	0	1	1
b5	1	0	0	1	1
b6	0	1	1	1	1

3.1. Result and Discussion

In this step we identify that how parameters are dispensable [1] so after that we work out to reduce the dimension of data for this task parameters can be removed without affected the original decisions [1].

Let us consider the representation of soft set (F, P) in tabular form. Suppose Q is a reduction of P , then the new soft set (F, Q) is called the reduct soft set of the soft set (F, P) [1].

Algorithm: Suppose Mr. XYZ selects bike:

The process of selecting a bike may has three basic steps

- Soft set theory can be used to transformed the dataset into a Boolean-valued information (as Table 1) system as

$S = (U, A, V(0, 1), f)$ soft set theory is used for this conversion [1].

- The next process is Input

Input the soft set (F, E) ,

Input the set P of choice parameters of Mr. XYZ which is a subset of E .

- At first the data set must be reduced by removing dispensable items before make a decision.

We used the soft set theory reduce parameters [1] [4].

Find all reduct-soft-sets of (F, P) , choose one reduct-soft-set say (F, Q) of (F, P) .

3.2. Explanation

It is known that $\{e1, e2, e4, e5\}$ and $\{e2, e3, e4, e5\}$ are two reducts of $P = \{e1, e2, e3, e4, e5\}$ [1]. But actually $\{e1, e2, e4, e5\}$ and $\{e2, e3, e4, e5\}$ are not the reducts of $P = \{e1, e2, e3, e4, e5\}$.

The following descriptions will explain this issue.

Let us consider R_p is the indiscernibility relation [1] [5] produced by $P = \{e1, e2, e3, e4, e5\}$, then the partition defined by R_p is $\{\{b1, b6\}, \{b3, b4, b5\}, \{b2\}\}$ based on Table-1.

The indiscernibility relation and the partition would be changed if one of parameter of $\{e1, e2, e3\}$ is deleted from P . so thus all of these three parameters are indispensable [1].

The partition would be changed to $\{\{b1, b6\}, \{b2, b3, b4\}, \{b2\}\}$, if $\{e1\}$ is removed from P .

The partition would be changed to $\{\{b1, b6\}, \{b3, b4, b5\}, \{b2\}\}$, If $\{e2\}$ is removed from P .

The partition would be changed to $\{\{b1, b6\}, \{b2, b3, b4, b5\}\}$, if $\{e3\}$ is removed from P .

The partition R_p would be unchanged, If $\{e4\}$ is removed from P .

The partition R_p would be unchanged, If $\{e5\}$ is removed from P .

So the conclusion is if $\{e4, e5\}$ is deleted from P , then the indiscernibility relation [1] and the partition R_p are not variant, so both of $e4$ and $e5$ are dispensable in P by Definition 2 [1].

So by Definition 3, $\{e1, e2, e3\}$ is the reduction of $P = \{e1, e2, e3, e4, e5\}$. From Table 1 it is proved that $e4$ and $e5$ are not relevant and is affect the choices of the bike since they take the same values for every bike.

So the Mr. XYZ is ready to make a decision to purchase bike based on the parameters $\{e1, e2, e3\}$.

We can apply Definition 3 to partition the objects based on the parameter co-occurrence and the support value [1].

For Table 1 data set, the following will be the co-occurrence:

Co-occurrence $(b1) = \{e2, e3, e4, e5\}$

Co-occurrence $(b2) = \{e1, e3, e4, e5\}$

Co-occurrence $(b3) = \{e1, e4, e5\}$

Co-occurrence $(b4) = \{e1, e4, e5\}$

Co-occurrence $(b5) = \{e1, e4, e5\}$

Co-occurrence $(b6) = \{e2, e3, e4, e5\}$.

The support value of each object can be given, by Definition 4, as

Support $(b1) = 4$ Support $(b2) = 4$

Support $(b3) = 3$ Support $(b4) = 3$

Support $(b5) = 3$ Support $(b6) = 4$

The data set can partitioned based on the support value of an object given by:

$\{\{b1, b2, b6\}, \{b3, b4, b5\}\}$

When the parameters $e4, e5$ are removed the current partitions would not be changed but for other parameter values it would be changed. So the conclusion is $e4$ and $e5$ are the reducts for the sample data set given in Table 1.

For the small sample set both methods produced the same reducts. So this task is important when dealing with a large data set. This method or algorithm is also beneficial to measures different types of performances of a data set. Feature selection is one of the important measures by using this method.

4. PERFORMANCE OF SOFT SET OVER OTHER APPROACHES

We can not use classical approaches to solve complicated problems in engineering, economics because of various uncertainties in these area. There are three theories which can consider as mathematical tools for dealing with uncertainties [2]. These theories are probability, fuzzyset and interval mathematics. Each theory has its own difficulty. Theory of probability can deal only with stochastically stable problems [2]. Interval mathematics deals the errors of calculations by constructing an interval estimate for the exact solution of a problem. This method is useful in many cases but this theory is not sufficiently adaptable for problem with different uncertainties [2]. Fuzzy set is ver powerful tool to deal with complicated problems but there exists a difficulty how to set the membership function in each particular case [2].

The concept of soft theory is a mathematical tool for dealing with uncertainties which is free from the above difficulties. It is a parameterized tool to deal with approximate value and uncertainty. Each approximate description can divide in two parts. First one is a predicate part and second one is an approximate value set [1].

5. CONCLUSION

In this paper, we discussed data uncertainty and non-crisp data. We used Soft set theory to convert a small data set to binary valued data set. We also discussed how to reduce the unnecessary parameters as reducts from the data set. The loss of these parameters does not affect the original information of the used data set. So Soft set theory is also used in dimensionality reduction and generally used to provide better and quick decision making in compare of previous theories.

REFERENCES

- [1] D. Ashok Kumar, R. Rengasamy, "Parameterization Reduction Using Soft Set Theory for Better Decision Making", Proceedings of the 2013 *International Conference on Pattern Recognition, Informatics and Mobile Engineering*, February 21-22.
- [2] D. Molodtsov, "Soft Set theory-first results", *Computer and Mathematics with Applications*.1999, 37 (4/5), 19-31.
- [3] Yao, Y.Y. "Relational interpretations of neighborhood operators and rough set approximation operators", *Information Sciences*, 1998,111, 239–259.
- [4] Tutut Herawan, Rozaida Ghazali, Mustafa Mat Deris, "Soft Set Theoretic Approach for Dimensionality Reduction", *International Journal of Database Theory and Application*. 2010, Vol. 3, No. 2.
- [5] P.K. Maji, A.R. Roy and R. Biswas, "An application of soft sets in a decision making problem". *Computer and Mathematics with Application*, 2002, 44(8/9), 1077-1083.
- [6] P.K. Maji, A.R. Roy and R. Biswas, "Soft Set Theory". *Computer and Mathematics with Application*, 2003; (45), 555-562,
- [7] Chen, D. "The parameterization reduction of soft sets and its applications", *Computers an Mathematics with Applications*, 2005,04/05.
- [8] Tutut Herawan. "Soft t Decision Making for Patients Suspected Influenza", *Lecture Notes in Computer Science*, 2010.
- [9] Manmath Kumar Bhuyan, Durga Prasad Mohapatra, Srinivas Sethi, "Software Reliability Prediction using Fuzzy Min-Max Algorithm and Recurrent Neural Network Approach", *Ineternational Journal of Electrical and computer Engineering (IJECE)*, 2016, 04/06.
- [10] Noor Cholis Basjaruddin, Kuspriyanto Kuspriyanto, Didin Saefudin, Ilham Khrisna Nugraha, "Developing Adaptive Cruise Control Based on Fuzzy Logic Using Hardware Simulation", *International Journal of Electrical and computer Engineering (IJECE)*, 2014, 04/06.
- [11] Janmejay Pant, Amit Juyal, Shivani Bahuguna, "Soft set, a soft Computing Approach for Dimensionality Reduction, *International Journal of Innovative Science, Engineering & Technology (IJSET)*, 2015, 04/02.

BIOGRAPHIES OF AUTHORS



He is working as an Asst. Professor in CSE and Applications department of Graphic Era Hill University; Bhimtal Campus. He completed his M.Tech in Information Technology from GEU Dehradun. He has an experience of around 07 years of teaching. His areas of interest are Machine Learning, Soft Computing and data Mining. He has taught several core subjects like C, JAVA, DBMS.OS etc. as well as advanced subjects like Soft Computing, Data Mining etc.



He received his M.Tech degree in computer science and engineering from Graphic Era University. His areas of interest are big data and data mining. He has published various papers in the different journals. Currently he is working as Assistant professor in Graphic Era Hill University Dehradun. He has an experience of around 09 years of teaching.



He received his MCA from UPTU Lucknow. His areas of interest are big data and Image Processing. He has published various papers in the different journals. Currently he is working as Assistant professor in Graphic Era Hill University Bhimtal. He has an experience of around 05 years of teaching



Mr. Akhilesh Dwivedi received the B.Tech in ETE from the MGM College of Engg & Techno, Noida (U.P. Technical Univ., Lucknow) India, in 2009 and M.Tech in Information Security from Ambedkar Institute of Technology, Govt. of NCT Delhi, Geeta Colony, New Delhi (GGSIU, New Delhi), India in 2011. His main research interests are in Big Data Analytics Data Mining, Biometric Security and Secure Semantic WebServices, Cryptography and Network Security, Data Storage Security in Cloud Computing. Mr. Dwivedi is the professional member of ISTE, CSTA, AIRCC, IAENG, IACSIT, and IAOE. He is the author/co-author of more than seventeen publications in International/National Journals and conferences.