International Journal of Scientific Research and Management (IJSRM)
||Volume||5||Issue||07||Pages||6552-6560||2017||
Website: www.ijsrm.in ISSN (e): 2321-3418
Index Copernicus value (2015): 57.47 DOI: 10.18535/ijsrm/v5i7.96

Interpolation Based Data Filling Approach of Incomplete Fuzzy Soft Set

Amalendu Si[1], Anup Kumar Kolya[2], Sudip Kumar Adhikari[3]

[1]Mallabhum Institute of Technology, Department of Computer Science and Engineering,
Bishnupur, Bankura 712402, India
[2]Department of Computer Science & Engineering,
RCC Institute of Information Technology, Kolkata, India
[3]Department of Computer Science & Engineering,
Cooch Behar Government Engineering College, West Bengal, India.

Abstract: ***Soft set theory which deals with vagueness and uncertainties are based on complete information. However, in practical problems, incomplete information also widely exists. The missing data in incomplete fuzzy soft set reduces the utilization of fuzzy soft set. In this paper, a new concept is introduced to estimate the missing data by the other incomplete fuzzy soft sets which improve the performance and importance of the incomplete soft set. The missing data is evaluated by the experts' observation using the interpolate relations among same properties and same object. An illustrative example shows that the accuracy and feasibility of estimated missing data is much better than the earlier reported techniques in a real application***

Keywords: Fuzzy soft Set, Incomplete Soft Set, Data Filling, Unknown Information.

1. Introduction

Soft set is a new concept and emerging mathematical tool proposed by Molodtsov in1999 [1]. Soft set deals with vagueness and it is subjective as well as objective uncertainty problem. Presently soft set theory is progressing rapidly with elementary properties and operations and it is introduced in different theoretical tools such as soft group, soft ordered semi-groups, soft ordered sub-semi group, soft rings and soft semi rings etc.

The soft group, introduced by H. Aktas and N. Cagman, is defined by the properties based on fuzzy soft set [2, 3]. Y. B. Jun, K. J. Lee and A. Khan reported the concept of soft ordered semi-groups in 2010. They investigated the various properties of soft ordered semi group and soft ordered sub-semi group [4]. The soft ring was modified into soft semi ring by Fenga, *et al.* in 2008 [5]. The soft semi group increases the utility in different areas of applied mathematics and information which provides an algebraic framework for modelling and investigating key factors in uncertainty problems [6]. Those tools overcome the limitation of inadequate parameterized tools used in classical methods such as theoretical probability, fuzzy set [7], vague set [8], interval soft set [9] and rough set theory [6,10]. Most of the researchers were interested on soft set concept after it was provided in different ways by Maji *et al.* [11] in 2001. Soft sets are modified differently depending on the situation and problem such as intuitionist fuzzy set [12], intuitionist fuzzy soft set [13], interval valued fuzzy soft set [14] and inter valued intuitionistic fuzzy soft set [15]. In a Parallel way, soft sets are applied in decision making [8, 16] and multi-criteria fuzzy decision making based on vague set theory. The membership degree of satisfiability and non-satisfiability of each alternative is considered based on a set of criteria which represents the ambiguous values. A useful way is provided that can help the decision makers to take decision efficiently. Z. Xiao, K. Gong and Y. Zou proposed the economic forecasting [17] combined with fuzzy soft set. The combined forecasting is better acceptance theory than the individual forecasting components. In medical diagnosis [18], a group of experts diagnose patients' diseases and also represent the patient status with the help of intuitionist multi fuzzy set. The different aggregation operators are applied to combine the experts' opinions and the information which identifies properties of diseases are compared. T. Herawan and M. M. Deris proposed a new technique named association rule mining [19] which sets the predetermined data items with uncertain and redundant data which improves the performance of rule mining.

Soft sets, mentioned earlier, are very applicable in our real life application. Socio economic environment like manufacturing, medical diagnosis, supplier selection or other decision making problems day to day become more complex. A single decision maker is not sufficient to take decision, but depends on number of experts according to complexity and importance of problem. It is simple, easy to implement and generates accurate outcome if the available data is complete. D. Tingquan *et al.* [20] introduced incomplete soft sets due to lack of experts, knowledge, experiences and resources. Some absence of data makes a set incomplete [21, 22, 23, 24] and therefore incomplete set makes unreliable and inaccurate outcome in decision making. This situation was handled by two approaches. Firstly, incomplete parts are removed from incomplete set and make it complete or estimate the incomplete part of the data set using average probability method [25] or the object parameter method proposed by Deng and Wang in [10] from available data. Qin *et al*. [26] proposed another method where the association degrees among the parameters are created and the highest priority is given to the parameter with highest degree.

Our paper analyzes the relations among the parameters of the incomplete fuzzy soft sets. A group of experts provide incomplete report where some data are missing based on their observations. The missing values of the parameters are not common; one expert may not justify one parameter which is considered by the other experts.

A parameter is missing in one data set but remaining data sets hold specific data values. Initially the missing data is considered as normal state and is assigned a membership value 0.5. Then an interpolation relation is created for the alternatives with missing value between same present parameters and missing parameters of each data set. With the help of this relations and existing data values the missing data can be estimated.

The soft set is assigned the choice value either 0 or 1 in but in fuzzy soft set the choice value is assigned by the membership value of the criteria. The associate degree is calculated depending upon the choice value which is not appropriate for finding the interior relations of an object. In this context, the distance between the two memberships values of parameters have been chosen to create the interior relations of an object. The authors [10] had not defined whether the missing value is more than one or not and Inconsistent Association Degree and consistent Association Degree are equal.

In our paper, the calculation of the relation between the parameters does not depend on the single pair. The calculation is done using all existing membership values of parameters which is provided by various experts. The weight of the parameter are calculated for estimating the missing values [10]. In this context, the calculation is done using difference between the membership values of the parameters.

The remaining part of the paper is explained as follows: Section 2 introduces fuzzy soft set. Interpolated fuzzy relation is explained in section 3. The proposed algorithm and illustrative example is explained in section 4 and section 5 respectively. Section 6 summarizes the conclusion.

2. Preliminaries

A pair (***F***, ***E***) is called fuzzy soft sets over ***U*** where ***F*** is a mapping given by $F: E \rightarrow f(U)$, a set of all fuzzy subsets in universe ***U***.

Let ***U***= $\{u_1, u_2 \ldots u_m\}$ be a set of *m* objects which may be characterized by a set of parameters ***E***= $\{e_1, e_2, e_3 \ldots\ldots\ldots e_n\}$. The fuzzy entity is a tabular representation of the fuzzy soft set $f(u_i, e_j)$ where, $f(u_i, e_j)$ is a quantity in the unit interval [0,1] which represents the degree of membership of the object u_i belonging to the parameter e_j or it can be defined as the membership degree or the membership probability of the object processing to the related parameter.

The fuzzy referential information must be collected before analyzing. During this collection four things have been considered which are nonempty finite set of alternatives (***U***), nonempty set of attribute (***A***), the domain value of the attribute ***A***, *v*[0,1] and an information function *f* specifying the attributes value for each object (defined by. $f: U \times A \rightarrow v$).

During the collection of information, some unknown data of the attributes appear, called incomplete system and also the unknown or missing data appear which is denoted by * in the tabular representation. Each fuzzy soft set can be considered as fuzzy referential information system in which each of the attributes (A) is a quantity which lies in the interval [0,1]. If the unknown value of the element in the fuzzy soft sets are there, then the fuzzy soft set can be considered as an incomplete information system.

Example 1. Let ***U*** be a set of the information system. The four objects are considered within the system given by ***U***= $\{u_1, u_2, u_3, u_4, u_5, u_6\}$, ***E*** is the collection of parameter set and parameters are in the word or sentence, ***E***= $\{e_1, e_2, e_3, e_4, e_5, e_6, e_7\}$ where e_1 and e_2 stand for 'beautiful' and 'size' respectively, e_3, e_4 and e_5 denote the 'cheap', 'in green surrounding' and 'modern' respectively. e_6 and e_7 indicate 'location and communication' and 'life time' respectively. The quality of the object is evaluated by the experts using four different properties. The experts cannot provide all the information of the properties in the object due to lack of knowledge and/or proper evaluation. This situation can be considered as an incomplete fuzzy soft set (***F***, ***E***).

The mapping ($\boldsymbol{F_i}$, ***E***) is defined as follows:

F(e_1) = {u_1/0.4, u_2/0.6, u_3/0.5, u_4/*, u_5/0.9, u_6/0.8}
F(e_2) = {u_1/*, u_2/0.7, u_3/0.8, u_4/0.4, u_5/0.5, u_6/0.7}
F(e_3) = {u_1/0.9, u_2/*, u_3/0.6, u_4/0.7, u_5/0.8, u_6/0.5}
F(e_4) = {u_1/0.7, u_2/0.6, u_3/*, u_4/0.8, u_5/0.4, u_6/0.6}
F(e_5) = {u_1/0.8, u_2/0.4, u_3/0.6, u_4/0.7, u_5/*, u_6/0.6}
F(e_6) = {u_1/0.6, u_2/7, u_3/0.8, u_4/*, u_5/0.6, u_6/0.7}
F(e_7) = {u_1/0.7, u_2/0.6, u_3/0.8, u_4/0.7, u_5/0.6, u_6/*}

Where, * indicates the unknown value of the parameters. The incomplete fuzzy soft set (***F***, ***E***) can be represented in the form of table.

Table 1: Tabular representation of the incomplete soft set (F, E)

u/e	e_1	e_2	e_3	e_4	e_5	e_6	e_7
u_1	0.4	*	0.9	0.7	0.8	0.6	0.7
u_2	0.6	0.7	*	0.6	0.4	0.7	0.6
u_3	0.5	0.8	0.6	*	0.6	0.8	0.8
u_4	*	0.4	0.7	0.8	0.7	*	0.7
u_5	0.9	0.5	0.8	0.4	*	0.6	0.6
u_6	0.8	0.7	0.5	0.6	0.6	0.7	*

3. Interpolate Attributes Relation

The computation of unknown points or values are tabulated using the surrounding points or values. Interpolation is the process of using known values $f(x_0)$, $f(x_1)$, $f(x_2)$,……, $f(x_n)$ to find the values for $f(y)$ at points $y \neq x_i$, where i=0,1,2….,n. In general, this technique involves the construction of a function $\boldsymbol{L}(x)$, called the interpolant. It agrees with f at the points $y=x_i$. Later it is used to compute the desired values. Extrapolation is the extension of data which beyond the range of the measurements. It is much more difficult, and it finds a serious errors if it is not used and interpreted properly. For example, a high-order polynomial may be fitted to the data set properly which over sit ranges of validity, but if higher powers need, it is included. The polynomial may be diverged rapidly from the smooth behavior of the data which belongs outside the range.

Fuzzy soft sets ($\boldsymbol{F_l}$, ***E***) are constructed based on the valuable fuzzy referential response, given by the experts. A set of experts ***D***= (d_1, d_2…. , d_p) are considered which are evaluated the same set of alternatives within the system ***U*** = {u_1, u_2, u_3………u_m}, with same set of attributers or criteria ***E*** = {e_1, e_2, e_3…………e_n}. Fuzzy entity of fuzzy soft sets is denoted by h_{ij}^l which indicates that the l-th expert gives the fuzzy reference of i^{th} alternative of j^{th} criteria. The unknown value called missing data in the incomplete fuzzy soft set is denoted by h_{ij}^l=" *" which indicates the l^{th} expert. The l^{th} expert does not provide any information of the i^{th} alternative of the j^{th} criteria. Incomplete fuzzy soft set ($\boldsymbol{F_l}$, ***E***) are constructed by l^{th} expert. The experts investigate the same set of alternatives with the same set of criteria. Hence, there is a dependence among the fuzzy quantity provided by the experts. One of the missing data in the attribute of the alternatives can be computed by the other relative data provided by the experts. The dependency of the relative data can be measured by the following method:

$$h_{ij}^l - h_{ij}^{l+1} \approx h_{ij+1}^l - h_{ij+1}^{l+1} \quad \text{or} \quad h_{ij}^l - h_{ij}^{l+1} \approx h_{i+1,j}^l - h_{i+1,j}^{l+1} \quad \text{……………………….…………. (1)}$$

The unknown data is computed with the help of above mentioned relation. Let h_{ij}^l is the unknown data and it is estimated with the following interpolate equation (2), where p number of experts are evaluated the alternatives with n number of criteria.

$$h_{ij}^l = \frac{1}{(p-1)(n-1)}\left(\sum_{k=1, k\neq l}^{p}\left(\sum_{b=1, b\neq j}^{n} h_{ij}^k\left(1-\left(\left|h_{ib}^l - h_{ib}^k\right|\right)\right)\right)\right) \qquad (2)$$

4. Algorithm

A new analytical method of incomplete fuzzy soft set is developed for filling the missing data of the experts' opinion using interpolate relation among the same properties.

i) A set of experts evaluated fuzzy referential information system with $\boldsymbol{U}=\{u_1, u_2, u_3, u_4, u_5, u_6\}$ be a set of alternatives and $\boldsymbol{E}=\{e_1, e_2, e_3, e_4, e_5, e_6, e_7\}$ be a universe set of parameters. Let $\boldsymbol{D}=\{d_1, d_2 \ldots d_p\}$ be the set of experts, d_l $(l=1,2,3\ldots\ldots p)$ gives the fuzzy reference entity for each alternatives, some of which are unknown due to lack of expertise or insufficient knowledge. According to the opinion of the experts, p number incomplete fuzzy soft sets can be constructed $(\boldsymbol{F_l}, \boldsymbol{E})$, l=1,2,3…..$p$.

ii) Find the unknown or missing data from the incomplete sift sets.

iii) The unknown entities of the incomplete fuzzy soft sets are considered as 0.5 during the evaluation process.

iv) Let h_{ij}^{l} is missing data of l^{th} expert of i^{th} alternative of j^{th} parameter. This missing data is calculated by the interpolation equation (Eq. 2).

v) If all missing data is filled up, stop the algorithm, otherwise go to step iii).

5. Illustrative Example

Suppose, four experts are conducting an evaluation using six alternatives of the information system $\boldsymbol{U}=\{u_1, u_2, u_3, u_4, u_5, u_6\}$ to find the better option. The set of attributes $\boldsymbol{E}=\{e_1, e_2, e_3, e_4, e_5, e_6, e_7\}$ represent the seven factors which are "beautiful", "size", "cheap", "in green surrounding", "modern", "communication and location" and "life time" respectively. But at the time of evaluation process some fuzzy reference quantity of the properties of the alternatives are unknown. As a result, the evaluation result is incomplete due to lack of expertise or insufficient experts' knowledge. According to the evaluation result of four experts we obtain the four incomplete fuzzy soft set (F_l, E), l=1,2,3,4.

After predicting the unknown values, the incomplete data sets are converted into complete data set. Next the average choice value of the alternatives which contain in the data sets of the attributes is calculated according to the best choice value.

Table 2: Tabular Representation of Incomplete Fuzzy Soft Set of Expert$_1$ $(\boldsymbol{F_1}, \boldsymbol{E})$.

u/e	e_1	e_2	e_3	e_4	e_5	e_6	e_7
u_1	0.4	*	0.9	0.7	0.7	0.6	0.4
u_2	0.6	0.7	*	0.6	0.8	0.7	0.5
u_3	0.5	0.8	0.6	0.5	0.6	*	0.7
u_4	*	0.4	0.7	0.8	0.5	0.7	0.8
u_5	0.6	0.8	0.5	0.7	*	0.5	0.6
u_6	0.5	0.7	0.9	0.6	0.7	0.9	*

Table 3: Tabular Representation of Incomplete Fuzzy Soft Set of Expert2 (F2, E).

u/e	e_1	e_2	e_3	e_4	e_5	e_6	e_7
u_1	0.6	0.5	0.8	*	0.5	0.7	0.8
u_2	*	0.8	0.6	0.8	0.5	*	0.7
u_3	0.4	*	0.7	0.4	0.8	0.9	0.6
u_4	0.7	0.7	*	0.6	0.7	0.7	0.7
u_5	0.6	0.8	0.9	0.5	*	0.6	0.7
u_6	0.8	0.6	0.8	0.7	0.8	0.7	*

Table 4: Tabular Representation of Incomplete Fuzzy Soft Set of Expert$_3$ $(\boldsymbol{F_3}, \boldsymbol{E})$.

u/e	e_1	e_2	e_3	e_4	e_5	e_6	e_7

u1	0.5	0.8	0.7	0.6	0.5	*	0.8
u2	0.8	*	0.5	0.7	0.6	0.8	0.6
u3	0.7	0.6	0.8	0.5	0.8	0.5	*
u4	0.6	0.5	0.6	*	0.9	0.6	0.7
u5	0.8	0.9	0.7	0.7	*	0.8	0.6
u6	*	0.6	0.6	0.8	0.7	0.6	0.5

Table 5: Tabular Representation of Incomplete Fuzzy Soft Set of Expert4 (F4, E).

Table 2 to Table 5 represent incomplete fuzzy soft sets of various experts. After observation we realize that all experts provide some of the missing data in their fuzzy soft set. There are twenty six missing data such as $F_1(u_1,e_2)$, $F_1(u_2,e_3)$, $F_1(u_3,e_6)$, $F_1(u_4,e_1)$, $F_1(u_5,e_5)$, $F_1(u_6,e_7)$, $F_2(u_1,e_4)$, $F_2(u_2,e_1)$, $F_2(u_2,e_6)$, $F_2(u_3,e_2)$, $F_2(u_4,e_3)$, $F_2(u_5,e_5)$, $F_2(u_6,e_7)$, $F_3(u_1,e_6)$, $F_3(u_2,e_2)$ $F_3(u_3,e_7)$, $F_3(u_4,e_4)$, $F_3(u_5,e_5)$, $F_3(u_6,e_1)$, $F_4(u_1,e_1)$, $F_4(u_2,e_4)$ $F_4(u_3,e_5)$, $F_4(u_4,e_6)$, $F_4(u_5,e_2)$, $F_4(u_5,e_7)$ and $F_4(u_6,e_2)$. Firstly, all missing data are filled by 0.5. The interpolate equation (Eq. 2) is applied to the attribute of e_2 of alternative u_1 in incomplete soft set of expert1. The missing value is

u/e	***e1***	***e2***	***e3***	***e4***	***e5***	***e6***	***e7***
u1	*	0.8	0.9	0.7	0.8	0.6	0.9
u2	0.6	0.7	0.8	*	0.6	0.5	0.7
u3	0.7	0.6	0.5	0.6	*	0.8	0.6
u4	0.5	0.4	0.7	0.8	0.5	*	0.8
u5	0.7	*	0.6	0.8	0.9	0.5	*
u6	0.8	0.5	*	0.7	0.6	0.8	0.9

estimated by the attribute value of the alternative u_1 in the other incomplete soft set.

From the general Eq. (1), putting l=1, i=1 and j=2 we get

$$h_{12}^{1}=\frac{1}{18}\left[\begin{array}{l}\left(\left(1-\left(\left|h_{11}^{1}-h_{11}^{2}\right|\right)\right)+\left(1-\left(\left|h_{13}^{1}-h_{13}^{2}\right|\right)\right)+\left(1-\left(\left|h_{14}^{1}-h_{14}^{2}\right|\right)\right)+\left(1-\left(\left|h_{15}^{1}-h_{15}^{2}\right|\right)\right)+\left(1-\left(\left|h_{16}^{1}-h_{16}^{2}\right|\right)\right)+\left(1-\left(\left|h_{17}^{1}-h_{17}^{2}\right|\right)\right)\right)*\mathrm{h}_{12}^{2}\\+\left(\left(1-\left(\left|h_{11}^{1}-h_{11}^{3}\right|\right)\right)+\left(1-\left(\left|h_{13}^{1}-h_{13}^{3}\right|\right)\right)+\left(1-\left(\left|h_{14}^{1}-h_{14}^{3}\right|\right)\right)+\left(1-\left(\left|h_{15}^{1}-h_{15}^{3}\right|\right)\right)+\left(1-\left(\left|h_{16}^{1}-h_{16}^{3}\right|\right)\right)+\left(1-\left(\left|h_{17}^{1}-h_{17}^{3}\right|\right)\right)\right)*\mathrm{h}_{12}^{3}\\+\left(\left(1-\left(\left|h_{11}^{1}-h_{11}^{4}\right|\right)\right)+\left(1-\left(\left|h_{13}^{1}-h_{13}^{4}\right|\right)\right)+\left(1-\left(\left|h_{14}^{1}-h_{14}^{4}\right|\right)\right)+\left(1-\left(\left|h_{15}^{1}-h_{15}^{4}\right|\right)\right)+\left(1-\left(\left|h_{16}^{1}-h_{16}^{4}\right|\right)\right)+\left(1-\left(\left|h_{17}^{1}-h_{17}^{4}\right|\right)\right)\right)*\mathrm{h}_{12}^{4}\end{array}\right] \quad (3)$$

From Eq. (3), we evaluate the h_{12}^{1} by putting the values of h_{11}^{1} to h_{17}^{1}, h_{11}^{2} to h_{17}^{2}, h_{11}^{3} to h_{17}^{3} and h_{11}^{4} to h_{17}^{4}.

$$h_{12}^{1}=\frac{1}{18}\left[\begin{array}{l}\left(\left(1-\left(\left|0.4-0.6\right|\right)\right)+\left(1-\left(\left|0.9-0.8\right|\right)\right)+\left(1-\left(\left|0.7-0.5\right|\right)\right)+\left(1-\left(0.7-0.5\right)\right)+\left(1-\left(0.6-0.7\right)\right)+\left(1-\left(0.4-0.8\right)\right)\right)*0.5\\+\left(\left(1-\left(\left|0.4-0.5\right|\right)\right)+\left(1-\left(\left|0.9-0.7\right|\right)\right)+\left(1-\left(\left|0.7-0.6\right|\right)\right)+\left(1-\left(0.7-0.5\right)\right)+\left(1-\left(0.6-0.5\right)\right)+\left(1-\left(0.4-0.8\right)\right)\right)*0.8\\+\left(\left(1-\left(\left|0.4-0.5\right|\right)\right)+\left(1-\left(\left|0.9-0.9\right|\right)\right)+\left(1-\left(\left|0.7-0.7\right|\right)\right)+\left(1-\left(0.7-0.8\right)\right)+\left(1-\left(0.6-0.6\right)\right)+\left(1-\left(0.4-0.9\right)\right)\right)*0.8\end{array}\right] \quad (4)$$

$$\frac{1}{18}\begin{bmatrix}(0.8+0.9+0.8+0.8+0.9+0.6)*0.5+\\(0.9+0.8+0.9+0.8+0.9+0.6)*0.8\\+(0.9+1+1+0.9+1+0.5)*0.8\end{bmatrix}$$

$$=\frac{1}{18}(4.8*0.5+4.9*0.8+5.3*0.8)$$

$$=\frac{1}{18}(2.4+3.52+4.24)=0.57$$

Similarly, other missing values are estimated as below like Eq. (3) and Eq. (4):

$h^1_{23}=0.57, h^1_{36}=0.71, h^1_{41}=0.51, h^1_{55}=0.57, h^1_{67}=0.54$, $h^2_{14}=0.53$, $h^2_{21}=0.67$, $h^2_{26}=0.53, h^2_{32}=0.47$, $h^2_{43}=0.56, h^2_{55}=0.59, h^2_{67}=0.55, h^3_{16}=0.57, h^3_{22}=0.62$, $h^3_{37}=0.54\, h^3_{44}=0.63$, $h^3_{55}=0.53\, h^3_{61}=0.61$, $h^4_{11}=0.46$, $h^4_{24}=0.63, h^4_{35}=0.6, h^4_{46}=0.58, h^4_{52}=0.58$ and $h^4_{63}=0.5$

Table 6 to Table 9 represent the corresponding complete fuzzy soft set of various experts and Table 10 depicts the choice value per alternative.

u/e	***e₁***	***e₂***	***e₃***	***e₄***	***e₅***	***e₆***	***e₇***
u₁	0.46	0.8	0.9	0.7	0.8	0.6	0.9
u₂	0.6	0.7	0.8	0.63	0.6	0.5	0.7
u₃	0.7	0.6	0.5	0.6	0.6	0.8	0.6
u₄	0.5	0.4	0.7	0.8	0.5	0.58	0.8
u₅	0.7	0.58	0.6	0.8	0.9	0.5	0.68
u₆	0.8	0.5	0.5	0.7	0.6	0.8	0.9

Table 6: Tabular Representation of complete Fuzzy Soft Set of Expert1 (***F₁***, ***E***).

u/e	***e₁***	***e₂***	***e₃***	***e₄***	***e₅***	***e₆***	***e₇***
u₁	0.4	0.57	0.9	0.7	0.7	0.6	0.4
u₂	0.6	0.7	0.5	0.6	0.8	0.7	0.5
u₃	0.5	0.8	0.6	0.5	0.6	0.71	0.7
u₄	0.51	0.4	0.7	0.8	0.5	0.7	0.8
u₅	0.6	0.8	0.5	0.7	0.57	0.5	0.6
u₆	0.5	0.7	0.9	0.6	0.7	0.9	0.54

Table 7: Tabular Representation of complete Fuzzy Soft Set of Expert2 (F2, E)

u/e	***e₁***	***e₂***	***e₃***	***e₄***	***e₅***	***e₆***	***e₇***
u₁	0.5	0.8	0.7	0.6	0.5	0.57	0.8
u₂	0.8	0.62	0.5	0.7	0.6	0.8	0.6
u₃	0.7	0.6	0.8	0.5	0.8	0.5	0.54
u₄	0.6	0.5	0.6	0.63	0.9	0.6	0.7
u₅	0.8	0.9	0.7	0.7	0.53	0.8	0.6
u₆	0.61	0.6	0.6	0.8	0.7	0.6	0.5

Table 8: Tabular Representation of Complete Fuzzy Soft Set of Expert3 (F3, E).

u/e	***e₁***	***e₂***	***e₃***	***e₄***	***e₅***	***e₆***	***e₇***
u₁	0.6	0.5	0.8	0.53	0.5	0.7	0.8
u₂	0.67	0.8	0.6	0.8	0.5	0.53	0.7
u₃	0.4	0.47	0.7	0.4	0.8	0.9	0.6
u₄	0.7	0.7	0.56	0.6	0.7	0.7	0.7
u₅	0.6	0.8	0.9	0.5	0.59	0.6	0.7
u₆	0.8	0.6	0.8	0.7	0.8	0.7	0.55

Table 9: Tabular Representation of Complete Fuzzy Soft Set of Expert4 (F4, E).

Table 10: Choice value per alternative

Alternative	***u₁***	***u₂***	***u₃***	***u₄***	***u₅***	***u₆***
Average Choice Value	0.66	0.65	0.63	0.65	0.67	0.68

6. Error Analysis

The absent values in the incomplete fuzzy set have been estimated based on available values of the incomplete fuzzy soft sets using interpolation. The estimated value is totally uncertain and actual value belongs to the range of 0<x<1. The estimated value may not be exactly equal to the actual absent value. There may be an error in the estimation because the calculation is based on indirect relation. If accuracy improves the error minimizes but never goes to zero. The accuracy can be defined as follows:

$$Accuracy = (1 - |Estimated\ value - Actual\ value|) * 100\ \% \quad (5)$$

Here, the range of the actual value and estimated value belongs to 0 and 1. When we get the estimated value as 1 on that time actual value is 0 and vise-versa and it fluctuates rapidly. Table 11 shows the error estimation of actual values and consider values.

Suppose, at any situation estimated value is 0 when actual value is 1 then the accuracy will be 0% [accuracy = (1-|0-1|) * 100=0%]. On the other hand, if we consider the absent actual value as 0.5 then accuracy is 50% [accuracy = (1-|0-0.5|) * 100=50%].

Table 11: Error estimation of actual values and consider values

Present value	*Actual Value*	*Estimated Value*	*Error = |Actual value - Estimated Value|*	*Consider Value*	*Error = |Actual value – Consider Value|*
*	0	0	0	0.5	0.5
*	0	1	1	0.5	0.5
*	1	0	1	0.5	0.5
*	1	1	0	0.5	0.5

'*' indicates the unknown value.

7. Result and Discussion

This section presents a comparative analysis of the proposed data filling approach from that of existing state of the art methods. The unknown values are filled by the consistence and inconsistence association degree among the parameters [26]. If the association degree is higher than the threshold value, the unknown value should be either 0 or 1 depending on other known parameter values, set by the experts. However, in fuzzy soft set, the data values are considered in the range of 0 to 1. Here, the known value is 0, 1 or any value in between 0 or 1. So, it is very difficult to measure the inconsistence and consistence association between parameters in fuzzy soft data set.

Zhu *et al.* [22] presented the method to fill the unknown data in the incomplete data set and also the authors created an association relation among the parameters in incomplete data set using equation $H(e_i,e_j) = \sum |F(e_i) - F(e_j)|$ and also calculated. The unknown values are estimated according to the both value of $H(e_i,e_j)$ and $H(e_i)$. Here the value of $H(e_i)$ considers only the closest parameter which fills the unknown data. The value of unknown data may be 0, 1 and an average value of the present data of the other parameters. But our proposed method fills the absent or unknown data values of the incomplete data sets after considering the known values of the corresponding parameters and criteria from all the incomplete data set. So, the proposed technique shows better consistent than the other existing methods.

Object parameter method [20] creates the relation between object and parameter which calculates the weight among them. In their work, the unknown values are estimated with the relation among object and parameter which belongs to same data sets, opinion of the expert.

According to [20] the weight considers $w_1=w_2=1/2$ by using the said procedure, the unknown values are calculated as $h_{24}=0.7363$, $h_{31}=0.5792$ $h_{34}=0.7403$, $h_{45}=0.3650$ and $h_{57}=0.4176$ also compared with average probability method which predicts the unknown values from the same data set. In this paper, we apply our proposed method on the above data set as well as others data sets and estimate the unknown value of the same data set which are $h_{24}=0.5592$, $h_{31}=0.40958$ $h_{34}=0.3452$, $h_{45}=0.5691$ and $h_{57}=0.39958$.

It is very difficult to say that which method is most appropriate because we do not know the value of the unknown parameter. All we are trying to predict the unknown value by separate way. The authors [20] use a single data set to analyze the forest fire problem and estimate the unknown value from the known values of the other parameters, which are very important at the time of the analysis. However, the set of experts of the group of multicriteria decision making method proposed their opinion in the form of incomplete dataset. The unknown data values are estimated with known values of the datasets. We have given the same importance of the experts' opinion and created a relation among them.

Therefore, we can state that according to the above results and discussion, our proposed technique shows the better performance than the others existing methods.

8. Conclusion

In this paper, we propose a data filling method to fill the missing data in the incomplete fuzzy soft sets. The experts observe a set of alternatives with same set of parameters and provide their reference to the soft sets with some missing data. Those missing data leads to wrong decision in the decision-making system or any other application of soft set. To fill those missing data, at first we create an interpolate relation among the available data provided by the experts. Then, with the help of other data set, part of the missing data and interpolated relation the missing data of incomplete data sets are estimated. We have verified the approach with an example with four expert opinions and our method provides satisfactorily results compared with other existing state of the art methods.

References

[1] D. Molodtsov, "Soft set theory First results", Computers and Mathematics with Applications, vol. 37, pp. 19–31, 1999.

[2] N. C. Agman, S. Enginoglu and F. C. Itak, "Fuzzy Soft Set Theory and Its Applications", Iranian Journal of Fuzzy Systems, vol. 8, no. 3, pp. 137-147, 2011.

[3] H. Aktas and N. Cagman, "Soft sets and soft groups". Information Sciences, vol. 177, pp. 2726–2735, 2007.

[4] Y. B. Jun, K. J. Lee and A. Khan, "Soft ordered semigroups", Math. Logic Quart., vol. 56, pp. 42–50, 2010.

[5] U. Acar, F. Koyuncu and B. Tanay, "Soft sets and soft rings", Computational Mathematical Application, vol. 59, pp. 3458–3463, 2010.

[6] F. Feng, Y. B. Jun and X. Zhao, "Soft Semi-rings". Computers and Mathematics with Applications, vol. 56, pp. 2621–2628, 2008.

[7] L. A. Zadeh, "Fuzzy sets, Inf. Control, vol. 8, pp. 338–356, 1965.

[8] S. M. Chen and J. M. Tan, "Handling multi criteria fuzzy decision-making problems based on vague set theory", Fuzzy Set Syst., vol. 67, pp. 163–172, 1994.

[9] Z. Xiaohong, "On Interval Soft Sets with Applications", International Journal of Computational Intelligence Systems, vol. 7, pp. 16-19, 2014.

[10] Y. Jiang, J. Wang, P. Deng and S. Tang, "Reasoning within expressive fuzzy rough description logics", Fuzzy Sets Syst., vol. 160, pp. 3403–3424, 2009.

[11] P. K. Maji, R. Biswas and A. R. Roy, "Intuitionistic fuzzy soft sets", Journal of. Fuzzy Math., vol. 9, pp. 677–692, 2001.

[12] K. Atanassov, "Intuitionistic fuzzy sets", Springer Physics, Heidelberg, 1999.

[13] Z. Li, G. Wen and Y. Han, "Decision making based on intuitionistic fuzzy soft sets and its algorithm", Journal of Computational Analysis and Applications, vol. 17, no. 4, pp. 620-631, 2014.

[14] X. B. Yang, T. Y. Lin, J. Y. Yang, Y. Li and D. J. Yu, "Combination of interval-valued fuzzy set and soft set," Computers & Mathematics with Applications, vol. 58, pp. 521-527, 2009.

[15] Y. Jiang, Y. Tang, Q. Chen, H. Liu and J. Tang, "Interval-valued intuitionistic fuzzy soft sets and their properties", Comput. Math. Appl., vol. 60, pp. 906–918, 2010.

[16] A. R. Roy and P. K. Maji, "A fuzzy soft set theoretic approach to decision making problems", J. Comput. Appl. Math., vol. 203, no. 2, pp. 412–418, 2007.

[17] Z. Xiao, K. Gong and Y. Zou, "A combined forecasting approach based on fuzzy soft sets", Journal of Computational and Applied Mathematics, vol. 228, pp. 326–333, 2009.

[18] S. Das, M. B. Kar and S. Kar, "Group multi-criteria decision making using intuitionistic multi-fuzzy sets", Journal of Uncertainty Analysis and Applications, pp. 1-10, 2013.

[19] T. Herawan, M. M. Deris, “A soft set approach for association rules mining”, Knowledge Based Syst., vol. 24, pp. 186–195, 2011.
[20] D. Tingquan and W. Xiaofei, “An object-parameter approach for predicting unknown data in incomplete fuzzy soft sets”, Applied Mathematical Modelling, vol. 37, pp. 4139–4146, 2013.
[21] Y. Zou and Z. Xiao, “Data analysis approaches of soft sets under incomplete information”, Knowledge-Based Syst., vol. 21, no. 8, pp. 941–945, 2008.
[22] D. Zhu, Y. Li, H. Nie and C. Huang, “A Decision-making Method Based on Soft Set under Incomplete Information”, IEEE 978-1-4673-5534-6, 2013.
[23] Minakshi, D. Rajan Vohra and Gimpy, “Missing Value Imputation in Multi Attribute Data Set”, International Journal of Computer Science and Information Technologies (IJCSIT), vol. 5, no.4, pp. 5315-5321, 2014.
[24] F. Feng, Y. B. Jun and X. Zhao: “Soft Semirings”. Computers and Mathematics with Applications, vol. 56, pp. 2621–2628, 2008.
[25] Y. Donga and Z. Xiao, “A Group Decision Making Method Based on Dempster-Shafer Fuzzy Soft Sets under Incomplete Information”, International Journal of Hybrid Information Technology, vol. 8, no. 3, pp. 287-296, 2015.
[26] H. Qin, X. Ma, T. Herawan, and J. M. Zain, “Data Filling Approach of Soft Sets under Incomplete Information”, ACIIDS 2011, LNAI, vol. 6592, pp. 302–311, 2011.

Author Profile

Amalendu Si received the B.Tech. degree in Information Technology form WBUT and M. Tech. degrees in Computer Science and Engineering from Kalyani Govt Engineering College in 2005 and 2007 respectively. Presently, he engages as a faculty member at Mallabhum Institute of Technology in Computer Science and Engineering department, WB, India from 2007.

Anup Kumar Kolya passed B. Tech in Computer Science & Technology from Kalyani University in 2002. He obtained the M.E. and Ph.D. degree in Computer Science & Engineering from Jadavpur University in 2008 and 2015 respectively. He had more than 5 years of research experiences and 5 years of teaching experience. His research interest includes Natural Language Processing, Fuzzy Logic and Bigdata. He is currently an assistant professor in Computer Science & Engineering Department in RCC Institute of Information Technology, Kolkata, India.

Sudip Kumar Adhikari passed B. Tech in Computer Science & Engineering from Vidyasagar University in 2002. He obtained the M.E. and Ph.D. degree in Computer Science & Engineering from Jadavpur University in 2008 and 2016 respectively. He had more than 13 years of teaching experiences. His research interest includes medical image processing, pattern recognition, artificial intelligence etc.. He is currently an assistant professor in Computer Science & Engineering Department in Cooch Behar Government Engineering College, West Bengal, India.