Songklanakarin J. Sci. Technol.
43 (4), 1063-1070, Jul. - Aug. 2021

http://www.sjst.psu.ac.th

Original Article

Dimensionality reduction - A soft set-theoretic and soft graph approach

Omdutt Sharma[1], Pratiksha Tiwari[2*], and Priti Gupta[3]

[1] *P. D. M. University, Bahadurgarh, Haryana, 124507 India*

[2] *Delhi Institute of Advanced Studies, Rohini, Delhi, 110085 India*

[3] *Maharashi Dayanand University, Rohtak, Haryana, 124507 India*

Received: 29 April 2020; Revised: 25 July 2020; Accepted: 12 August 2020

Abstract

Due to the digitization of information, organizations have abundant data in databases. Large-scale data are equally important and complex hence gathering, storing, understanding, and analyzing data is a problem for organizations. To extract information from this superfluous data, the need for dimensionality reduction increases. Soft set theory has been efficaciously applied and solved problems of dimensionality, which saves the cost of computation, reduces noise, and redundancy in data. Different methods and measures are developed by researchers for the reduction of dimensions, in which some are probabilistic, and some are non-probabilistic. In this paper, a non-probabilistic approach is developed by using soft set theory for dimensionality reduction. Further, an algorithm of dimensionality reduction through bipartite graphs is also described. Lastly, the proposed algorithm is applied to a case study, and a comparison of results indicates that the proposed algorithm is better than the existing algorithms.

Keywords: dimensionality reduction, soft set, grade membership, binary-valued information system, soft graph

1. Introduction

Due to digitalization, there is a rapid growth in the amount of information/data and the effects of this abundance lead to difficulty in managing information, which can lead to an overload of data that contains irrelevant and redundant data. Handling such problems can ravage plenty of time and money. To deal with these types of problem, there is a need to eliminate irrelevant redundant data by a technique which is known as dimensionality reduction. Dimensionality reduction has been a prolific topic of study and growth, since the last four to five decades. It has been exceptionally beneficial in eliminating avoidable and repetitive data, increasingly effectiveness in various areas (Gupta & Sharma, 2015). Dimension reduction can be useful in reducing cost, redundancy, and noise. Thus, it is one of the best tools to deal with real-life problems.

*Corresponding author
Email address: parth12003@yahoo.co.in

Numerous real-life situations consist of uncertainty that cannot be successfully modeled by the classical mathematical theories. To handle such problems contemporary mathematical/ statistical theories were developed. Probability theory (Kolmogorov, 1933), fuzzy sets theory (Zadeh, 1965), rough sets theory (Pawlak, 1982), intuitionistic fuzzy sets theory (Atansassov, 1986) and vague sets theory (Gau & Buehre, 1993) are some of the key notions. Molodtsov (1999) pointed out the different limitations of some of these theories. The reason for these limitations is possibly the inadequacy of the parameterization tool of the theories, and consequently Molodtsov (1999) presented the notion of soft theory as a novel mathematical tool which overcomes these limitations and successfully applied it to the theory of games, measurement theory, smoothness of function and Riemann integration.

Dimensionality reduction can help to solve various decision-making problems by reducing attributes of the original data using soft set theory. As far as the standard soft sets are concerned, it can be defined as two different equivalence classes of objects, thus confirming that Boolean-valued information valued systems can be dealt with by soft sets.

Maji, Roy and Biswas (2002), defined algebraic operations for soft sets and verified binary operations for the same. Chen, Tsang, Yeung and Wang (2005), improved the method anticipated by Maji *et al.* (2002). The fuzzy soft theoretic approach was developed to solve decision-making problems by Roy and Maji (2007). Zhao, Luo, Wong and Yao (2007) presented definition of reduct which jointly sufficient and individually necessary for preserving the properties of a given information table. Kong, Gaoa, Wang and Li (2008) derived a heuristic normal method in soft and fuzzy soft sets for parameter reduction. Herawan, Rose and Mat Deris (2009) developed an approach for reduction of attributes in a multi-valued information system in a soft theoretic environment, which is equivalent to Pawlak's rough reduction. Alcantud (2016) examined relationships amongst the soft sets and other theories. Zhan and Alcantud (2017) reviewed soft and fuzzy soft based algorithms for parameter reduction.

Different mathematical theories focus on respective aspects such as fuzzy, rough set and soft set theories focusing on membership degree, granular and parameterization respectively. Many researchers have worked on the soft set theory which has a parameterization tool, but certain situations involve non-Boolean datasets that require hybridization. Thus, authors developed hybrid theories such as fuzzy soft sets (Agman, Enginoglu, & Citak, 2011), soft rough sets (Feng *et al.*, 2011), fuzzy rough and rough fuzzy sets (Dubois & Prade, 1990), Soft rough fuzzy(SRF) and soft fuzzy rough (SFR) sets (Meng *et al.*, 2011), etc. to deal with such type of data. Zhang and Wang (2018) investigated types of soft coverings based on rough sets and their properties. Ma, Zhan, Ali & Mehmood (2018) reviewed decision making methods based on hybrid SRF and SFR sets. Zhang, Zhan and Alcantud (2019) concepts of fuzzy soft β-minimal and β-maximal descriptions their β-coverings and relationships amongst them. Zhang, Zhan and Yao (2020) presented covering-based variable precision intuitionistic fuzzy rough set models and applied multi-attribute decision-making problems of bone transplants. Similarly, the two papers of Jiang, Zang, Sun and Alcantud (2020) and Ma, Zang, Sun and Alcantud (2020) worked on covering-based variable precision fuzzy rough sets and multi-granulation fuzzy rough set respectively and applied them to multi-attribute decision making.

There is another non-probabilistic approach known as graph theory, helps to model various real-life situations. This theory is a suitable tool for solving combinatorial problems in different areas such as geometry, algebra, number theory, topology, operation research, optimization, and computer science, etc. Researchers Rosenfeld (1975), Thumbakara and George (2014), and Mohinta and Samanta (2015) combined graph theory with the fuzzy, soft set, and fuzzy soft set theory, respectively. Smarandache, (2018) generalized soft set to hypersoft set by transforming the function F into a multi-attribute function. Thereafter, introduced the hybridization of Crisp, Fuzzy, Intuitionistic Fuzzy, Neutrosophic, and Plithogenic Hypersoft Set. Classical techniques for reduction of dimension are principal component analysis and multi-dimension scaling; both are based on the concept of distance. Graphs provide an effective way to encrypt neighborhood relations. Burianek, Zaoralek, Snasel, and Peterek (2015) studied select dimension reduction techniques and used them to draw sensible graphs from initial graphs, and compared techniques with the technique of Kamada and Kawai (1989). Jatram and Biswas (2015) proposed a multiple dimension reduction method of feature space of graphs by using Spectral methods for the FANNY clustering algorithm. Qian, Yin, Kong, Wang, and Gao (2019) presented an algorithm for low-rank graph optimization for Multi-View Dimensionality Reduction.

Based on the soft theoretic approach, here, a non-probabilistic approach in dimensionality reduction is developed. Then Boolean-valued information system is represented as a bipartite graph and an algorithm is presented that can be used for dimensionality reduction.

2. Related Concepts of Soft Sets and Graph Theory

A parameterized mathematical tool; soft set theory deals with a collection of objects with categories defined approximately. Each category has two sections - a predicate and an estimated/ approximate value set. Since the initial portrayal of the object has an approximate nature, the notation of exact solution is also not required. The non-appearance of any confinements on the approximate depictions, in soft set theory, makes it entirely appropriate and just material practically speaking. With the assistance of words and sentences, real number, function, mapping, etc. any parameterization can be utilized. Hence predicament of the membership function or any related issue does not exist in the soft set theory.

Definition 1 (Soft Sets): Let U represent the initial universe set and P represent a set of parameters. Then the ordered set (m, P) is known as soft set (on the initial universe, U) if m is a mapping defined from P into Pow(U), where Pow(U) represents the power set of U, i.e.

m: P $\rightarrow$ Pow(U).

Clearly, (m, P) on initial universe U represents the family of subsets parameterized over U. Also, m(p) represents a set of p-approximate element, for any p ϵ P for the soft set (m, P).

Example 1.1: Let U = {b_1, b_2, …….., b_5} represents different model of bikes and P represents different selection criteria/ parameters, P = {p_1 = expensive, p_2 = beautiful, p_3 = cheap, p_4 = in good repair, p_5 = good mileage}. Then the attractiveness of the bikes can be represented by (m, P) i.e. soft set.

Example 1.2: Let F and m_F represent a fuzzy set and its membership function respectively i.e., $m_F : U \rightarrow [0, 1]$. Let P(α) = {x ϵ U: $m_F(x)$ greater than equal to α}, where α ϵ [0, 1] represents family of α- level sets for m_F. $m_F(x)$ can be defined as $m_F(x) = \underset{x \epsilon P(\alpha)}{sup_{\alpha \epsilon [0.1]}} \alpha$, where family P is known.
Thus, every fuzzy set F can be represented as a soft set.

Alternative definition of soft sets: A pair (M, U) is defined as a soft set over P where M is a function defined from initial universe U to Power set of P. Alternatively in other words, a soft set is a family of subsets of the universal parameter sets of P. Here U is the set of objects and P is the universal set of parameters, where parameters can be properties/ characteristics/attributes of objects.

Example 1.3: Let U = {b_1, b_2,…….., b_5} be the different sets of cars under consideration and P be the set of parameters, P = {p_1 = price, p_2 =looks, p_3 = speed, p_4 = weight, p_5 = average}. Then the soft set (M, U) describes the attractive cars.

2.1 Information systems equivalence with soft sets

Definition 2: Let a finite set of objects and attributes be denoted by U and A respectively. Then the quadruple (U, A, V, g) is known as an information system, where $g : U \times A \to V$ i known as information function and $V = \cup_{a\in A} V_a$. Also, the value of attribute a is given by $V_a = \{g(x, a) | a \in A \text{ and } x \in U\}$.

Alternatively, An information system is known as a knowledge representation/attribute-valued system and can be spontaneously represented in the form of an information table and it reduces to Boolean-valued information system if $V_a = \{0,1\}$, for every $a \in A$, in information system $S = (U, A, V, g)$.

Proposition 1: Let (F, E) be any soft set over initial universe U, then soft set (F, E) can be represented as Boolean-valued information system $S = (U, A, V_{\{0,1\}}, g)$ and vice versa.

Proof: Consider (F, E) and define a mapping

$F = \{g_1, g_2, \ldots \ldots \ldots \ldots, g_n\}$.

Where

$$g_1 : U \to V_1 \;\; and \;\; g_1(x) = \begin{cases} 1, & x \in F(e_1) \\ 0, & x \notin F(e_1) \end{cases}$$

$$g_2 : U \to V_2 \;\; and \;\; g_2(x) = \begin{cases} 1, & x \in F(e_2) \\ 0, & x \notin F(e_2) \end{cases}$$

……………………………………………….

………………………………………………..

$$g_n : U \to V_n \;\; and \;\; g_n(x) = \begin{cases} 1, & x \in F(e_n) \\ 0, & x \notin F(e_n) \end{cases}$$

Thus, for $A = E$ and $V = \cup_{e_i\in A} V_{e_i}$, for $V_{e_i} = \{0, 1\}$, then any soft set (F, E) can be considered as a $S = (U, A, V_{\{0,1\}}, g)$ which is a Boolean-valued information system and vice versa.

Definition 3: Let $S = (U, A, V_{\{0,1\}}, g)$ be a binary-value information system. Thus (F_S, A) is called the soft set over U induced by S, where $F_S : A \to 2^U$ and for any $x \in U$ and $a \in A$, $F_S(a) = \{x \in U \mid g(x, a) = 1 \text{ or } 0\}$.

Hence, the soft set given in Example 1.1 can be denoted as a Boolean-valued information system represented in Table 1.

Example 1.4: In example 1.1, U and P represents the set of objects and attributes/ parameters respectively, and we consider $V = \{0,1\}$ to be the sets of values of those objects which satisfy the parametric conditions. Here {1} and {0} represent that condition is satisfied and not satisfied respectively. Then we define

$g(b_1, p_1) = 1,\quad g(b_1, p_2) = 1,\quad g(b_1, p_3) = 0,\quad g(b_1, p_4) = 1,\quad g(b_1, p_5) = 0$

$g(b_2, p_1) = 1,\quad g(b_2, p_2) = 0,\quad g(b_2, p_3) = 0,\quad g(b_2, p_4) = 1,\quad g(b_2, p_5) = 0$

$g(b_3, p_1) = 1,\quad g(b_3, p_2) = 1,\quad g(b_3, p_3) = 1,\quad g(b_3, p_4) = 1,\quad g(b_3, p_5) = 1$

$g(b_4, p_1) = 0,\quad g(b_4, p_2) = 1,\quad g(b_4, p_3) = 1,\quad g(b_4, p_4) = 1,\quad g(b_4, p_5) = 1$

$g(b_5, p_1) = 1,\quad g(b_5, p_2) = 0,\quad g(b_5, p_3) = 0,\quad g(b_5, p_4) = 1,\quad g(b_5, p_5) = 1$

2.2 Soft set: Tabular representation

Lin (1998) and Yao (1998) represented soft sets in Tabular form. This section signifies analogues representation in binary table. Consider the soft set (m, P) for the set of parameters P. Soft set in a tabular form can be represented as binary table given in Table 1. Binary representation is beneficial for storing soft set in computer memory. The tabular representation of example1.4 is as

Table 1**.** Binary representation

U/P	p_1	p_2	p_3	p_4	p_5
b_1	1	1	0	1	0
b_2	1	0	0	1	0
b_3	1	1	1	1	1
b_4	0	1	1	1	1
b_5	1	0	0	1	1

In Table 1, if $b_{ij} \in m(p_i)$ then $b_{ij} = 1$, otherwise $b_{ij} = 0$, where b_{ij} are the entries in its table. Here entries are in the form of 0 and 1 thus it is known as a Boolean-valued information system table.

2.3 Graph theory: Some basic concepts

A graph G is represented by an ordered pair (V, E), where V is consists of a non-empty set of objects called vertices and E is a set of relation defined between two elements of V called edges. Two vertices x and y are said to be adjacent if $\{x, y\} \in E$. Subgraph of a graph G is denoted by $G' = (V', E')$ where $V' \subseteq V$ and $E' \subseteq E$. Two vertices are adjacent in subgraph G' if and only if they are adjacent to G. If all the vertices of G are connected to every other vertex then it is known as complete graph and represented as K_n where n is number of vertices.

Definition 4: A graph G is known as bipartite graph if vertex set of G can be partitioned in to two sets $V_1$and V_2such that $V(G) = V_1 \cup V_2$ and $V_1 \cap V_2 = \emptyset$. A complete bipartite graph is a bipartite graph if there exists a unique edge between each vertex of $V_1$and every vertex of V_2.

2.4 Soft graph:

Soft sets can deal with various types of uncertainties mathematically. Pictorial representation of soft sets is represented by soft graphs. Every simple graph can be represented as a soft graph whereas in this section we prove that every soft set can be represented as a bipartite graph. Following are some basic concepts:

Definition 5: A graph G represented by quadruple (G^*, F, K, A) is said to be a soft graph if following axioms are satisfied:

a) $G^*(V, E)$ represents a simple graph.
b) Set A represents a non-empty set of parameters.
c) (F, A) and (K, A) both represent soft set on V and E respectively.
d) For all $a \in A$,$(F(a), K(a))$ represents subgraph of G^*.

The collection of all subgraphs of G is represented by $SG(G)$.

Example: Consider the graph $G = (V, E)$ as shown in Figure 1. Let $A = \{1,5\}$. Define the set valued function F by, $F(x) = \{y \in V | x\ R\ y \Leftrightarrow d(x, y) \leq 2\}$.

Then $F(1) = \{1,2,3\}, F(5) = \{3,4,5\}$. Here $F(x)$ is a connected subgraph of G, for all $x \in A$. Hence $(F, A) \in SG(G)$.

Definition 5: Bipartite soft graph is defined as $V(G) = V_1 \cup V_2$ where $V_1 \cap V_2 = \emptyset$, where V_1 and V_2 represent set of parameters and objects respectively such that every edge of G joins a vertex of V_1 to a vertex of V_2.

Every simple graph can be represented as a soft set and every soft set can be represented as a bipartite graph (Hussain *et al.*, 2016).

3. The Proposed Techniques

This section introduces a novel concept for the reduction of parameters and objects (Dimensionality reduction). Here the idea is to reduce dimensions of data without changing the decision.

Definition 7: Let (m, E) represents a soft set, and m is defined as $m : E \rightarrow Pow(U)$; where E and U represent parameters and universal set of objects respectively.

Let $E = \{e_1, e_2, e_3, e_4\}$ and U = $\{p_1, p_2, p_3, p_4\}$ then $m_E(e_1)$ =$\{p_1, p_2, p_3\}$, $m_E(e_2)$= $\{p_3, p_4\}$, $m_E(e_3)$ = $\{p_1, p_2, p_3, p_4\}$, $m_E(e_4)$= $\{p_1, p_3, p_4\}$, then we define a measure:

$$\gamma_E(e_i) = \frac{card(m_E(e_i))}{card(U)} \qquad \text{Where } 0 \leq \gamma_E(e_i) \leq 1 \text{ for all } i;$$

where card(X) represents number of elements in set X.

Every set $m_E(e_i)$ for $e_i \in E$ from the parameterized family of subsets of the set U may be considered as the set of e_i-elements of the soft set (m, E) or as the set of e_i-approximate elements of the soft sets. $\gamma_E(e_i)$ is the grade of membership of e_i in universal set U. Here $\gamma_E(e_1) = 3/4$, $\gamma_E(e_2) = 1/2$, $\gamma_E(e_3) = 1$ and $\gamma_E(e_4) = 3/4$.

Definition 8: Let (M, U) is parameterized valued soft sets then $M : U \rightarrow Pow(E)$; where $Pow(E)$ represents power set of universal parameterized set E and U is set of objects.

Let E = $\{e_1, e_2, e_3, e_4\}$ and U = $\{p_1, p_2, p_3, p_4\}$ then $M_U(p_1)$ = $\{e_1, e_3, e_4\}$, $M_U(p_2)$ = $\{e_1, e_3\}$, M_U (p_3)= $\{e_1, e_2, e_3, e_4\}$ and M_U (p_4)= $\{e_2, e_3, e_4\}$ then we define a measure:

$$\sigma_U(p_i) = \frac{card(M_U(p_i))}{card(E)};$$

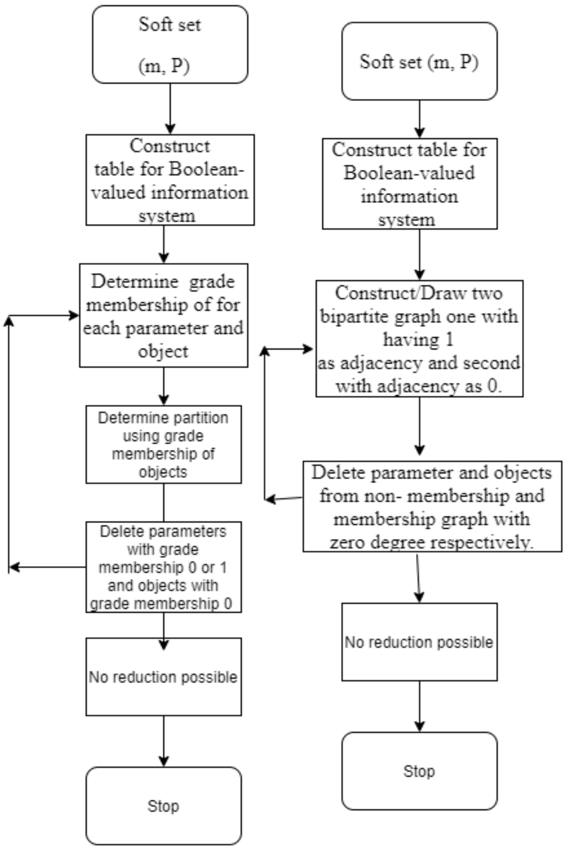

Figure1(i) Figure1(ii)

Figure 1. Flowchart for algorithms

where$0 \leq \sigma_U(p_i) \leq 1$ for all i; where $card(E)$ represents the number of elements in E.

Every set $M_U(p_i)$ for $p_i \in U$ is the subsets of universal parametrized set E or as the set of p_i-approximate elements of the parameterized valued soft sets. $\sigma_U(p_i)$, is the grade of membership of p_i in the parameterized universal set E. Here $\sigma_U(p_1) = 3/4$, $\sigma_U(p_2) = 1/2$, $\sigma_U(p_3) = 1$, $\sigma_U(p_4) = 3/4$.

3.1 Algorithm for dimensionality reduction by soft set technique

Input: The soft set (m, P)

(i) Construct table for Boolean-valued information system with the help of soft set (m, P).
(ii) Determine $\gamma_E(e_i)$ and $\sigma_U(p_j)$.
(iii) Determine the cluster partition U/E according to the value of $\sigma_U(p_i)$.
(iv) Delete those parameters and objects for which $\gamma_E(e_i) = 0$ *and* 1 and $\sigma_U(p_j) = 0$ respectively.
(v) Now for reduced parameters and objects go to step (ii) and repeat the process.
(vi) If there is no reduction possible then the Boolean-valued information system table is our desired dimensionality reduced table.

Figure 1(i) represents flowcharts for the above-mentioned algorithm.

Output: The dimensionality reduced the Boolean-valued information table which gives information in decision making. Remark: Assume that the number of objects and attributes in the fuzzy soft set (m, P) be n and m respectively. For calculating $\gamma_E(e_i)$ and calculating $\sigma_U(p_j)$ comparing each entry the complexity of computing the table is $O(n^2)$.

3.2 Dimensionality reduction using soft Set

In this section example presented by Maji *et al.* (2002) analyzed which was also discussed by Chen *et al.* (2005). "Let U = $\{h_1, h_2, h_3, h_4, h_5, h_6\}$ be a set of six houses, E = {expensive, beautiful, wooden, cheap, in green surroundings, modern, in good repair, in bad repair} be the set of parameters. Let Mr. X is interested to buy a house on the subset of the following parameter P = {beautiful, wooden, cheap, in green surroundings, in good repair}."

Consider the set of parameters P represented by $\{p_1, p_2, p_3, p_4, p_5\}$ symbolically. Boolean-valued information system table gives the soft set as in Table 2(a). Now determine $\gamma_P(p_i)$ and $\sigma_U(h_i)$ using Table 2(a) given in Table 2(b). As $\gamma_P(p_1)$ and $\gamma_P(p_3)$ both are equal to 1, thus remove p_1 and p_3 thus Table 2(b) reduces to Table 2(c). In Table 2(c) $\sigma_U(h_5) = 0$, thus remove h_5, now Table 2(c) reduces to Table 2(d). Again Table 2(d), $\gamma_P(p_4) = 1$, remove p_4, now it reduced to Table 2(e), again $\sigma_U(h_4)$=0, remove h_4 which reduces to Table 2(f). Clearly, further reduction of dimensionality is not possible, so this is the desired reduction. Here the proposed algorithm can eliminate more parameters without changing the decision parameter described by Maji *et al.* (2002). Thus, proposed technique is better than that of Maji *et al.* (2002).

3.3 Reduction of parameter using bipartite graphs

Each Boolean information soft set can be characterized by a bipartite graph. Each finite set of objects and set of parameters represented by two vertex sets of bipartite graph and list of adjacency contains all the 1's. To reduce dimensionality, the algorithm is as follows:

3.4 Algorithm for dimensionality reduction by soft set technique

Input: The soft set (m, P)

(i) Construct table for Boolean-valued information system by using a soft set (m, P).
(ii) Construct two bipartite graph one with having 1 as adjacency and second with adjacency as 0. (where one vertex set is a set of objects and the other vertex set is a set of parameters) and name them as membership and non- membership graphs.
(iii) Delete parameter from non- membership graph with zero degree.
(iv) Redraw the membership graph.
(v) Delete objects with degree zero from the membership graph.
(vi) Redraw both membership and non-membership graphs go to step (iii) and repeat the process until no reductions possible.

Figure 1**(**ii**)** represents flowcharts for the above-mentioned algorithm**.**

Table 2**.** Dimensionality reduction

Table 2(a)	U/P	p_1	p_2	p_3	p_4	p_5
	h_1	1	1	1	1	1
	h_2	1	1	1	1	0
	h_3	1	0	1	1	1
	h_4	1	0	1	1	0
	h_5	1	0	1	0	0
	h_6	1	1	1	1	1

Table 2(b)	U/P	p_1	p_2	p_3	p_4	p_5	$\sigma_U(h_i)$
	h_1	1	1	1	1	1	1
	h_6	1	1	1	1	1	1
	h_2	1	1	1	1	0	4/5
	h_3	1	0	1	1	1	4/5
	h_4	1	0	1	1	0	3/5
	h_5	1	0	1	0	0	2/5
	$\gamma_P(p_i)$	1	½	1	5/6	½	

Table 2(c)	U/P	p_1	p_2	p_3	p_4	p_5	$\sigma_U(h_i)$
	h_1		1		1	1	1
	h_6		1		1	1	1
	h_2		1		1	0	2/3
	h_3		0		1	1	2/3
	h_4		0		1	0	1/3
	h_5		0		0	0	0
	$\gamma_P(p_i)$		1/2		5/6	½	

Table 2(d)	U/P	p_1	p_2	p_3	p_4	p_5	$\sigma_U(h_i)$
	h_1		1		1	1	1
	h_6		1		1	1	1
	h_2		1		1	0	2/3
	h_3		0		1	1	2/3
	h_4		0		1	0	1/3
	h_5						
	$\gamma_P(p_i)$		3/5		1	3/5	

Table 2(e)	U/P	p_1	p_2	p_3	p_4	p_5	$\sigma_U(h_i)$
	h_1		1			1	1
	h_6		1			1	1
	h_2		1			0	½
	h_3		0			1	½
	h_4		0			0	0
	h_5						
	$\gamma_P(p_i)$		3/5			3/5	

Table 2(f)	U/P	p_1	p_2	p_3	p_4	p_5	$\sigma_U(h_i)$
	h_1		1			1	1
	h_6		1			1	1
	h_2		1			0	½
	h_3		0			1	½
	h_4						
	h_5						
	$\gamma_P(p_i)$		3/4			3/4	

Output: Dimensionality-reduced graph which gives information in decision making

Remark: Assume that the number of objects and attributes in the fuzzy soft set (m, P) be n and m respectively. For drawing both membership and non-membership graph the complexity is $O(n)$. Since graphs are redrawn for maximum of 'm' attributes, resulting in the complexity of algorithm as $O(n^2)$.

3.5 Dimensionality reduction using bipartite graph

This section discusses the example in section 3.2 using bipartite. Let U = {House 1(h_1), House 2(h_2), House 3(h_3), House 4(h_4), House 5(h_5), and House 6(h_6)} be a set of six houses, Let Mr. X is interested to buy a house on the following parameters subset P = {beautiful (P_1 i.e. 7), wooden (P_2 i.e. 8), cheap (P_3 i.e. 9), in green surroundings (P4 i.e. 10), in good repair (P_5 i.e. 11)}. Let {7,8,9,10,11} graphically represents parameters.

Consider the set of parameters and objects into the two disjoint sets of vertices of bipartite graph and edges shows the relationship between them. By using Table 2, we draw two types of bipartite graph as represented in Figure 2 (i) membership graph and Figure 2(ii) non-membership graph below:

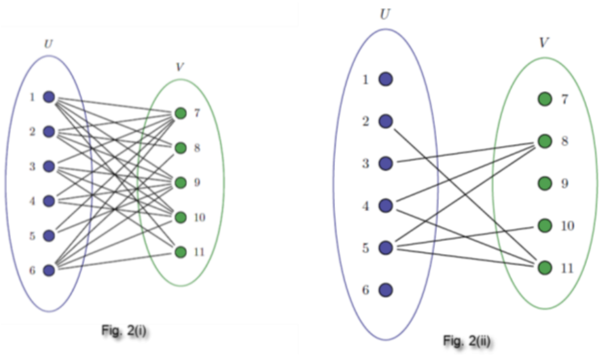

Figure 2. (i) Membership graph (ii)Non-membership graph

Here Figure 2(i) represents sets of those objects which satisfy the parametric conditions and Figure 2(ii) represents sets of those objects which does not satisfy the related parametric condition. From Figure 2 (ii) we see that degree of 7 and 9 is zero, we remove these parameters and again draw the bipartite graph shown in Figure 3 reduced membership graph.

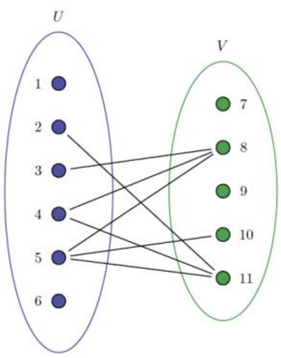

Figure 3. Reduced membership graph

From Figure 3 we see that house 5 has degree zero, so delete house 5 and redraw membership and non-membership bipartite graphs Figures 4(i) and 4(ii) respectively.

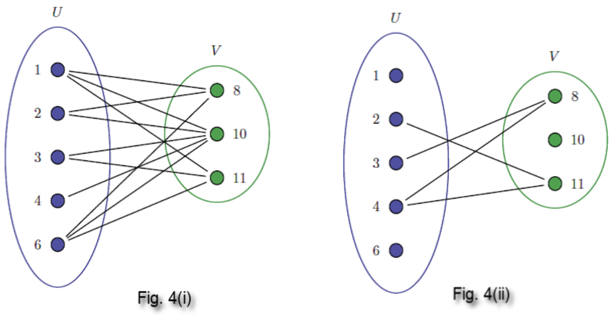

Figure 4. (i) Reduced membership graph (ii) Reduced non-membership graph

From reduced non-membership graph Figure 4(ii) we see that parameter 10 has degree zero. Thus parameter 10 can be deleted from the graph and the further reduced membership graph is represented by Figure 5.

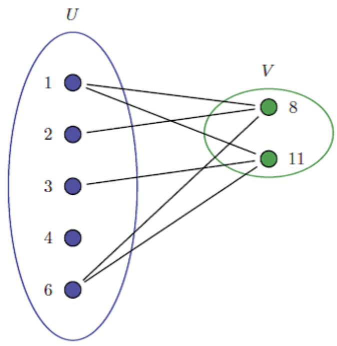

Figure 5. Reduced membership graph

Again, from Figure 5 it is observed that house 4 has degree zero. Again, redraw both membership and non-membership graph represented by Figures 6(i) and 6(ii).

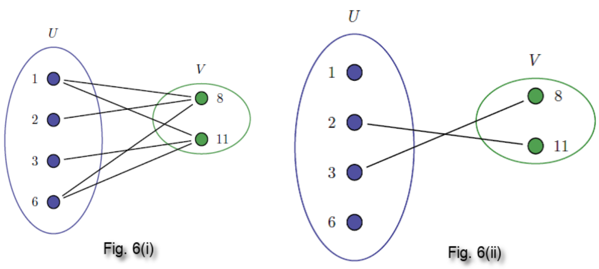

Figure 6. (i) and (ii): Final reduced graph

Here Figure 6(i) represents sets of those objects which satisfy the parametric conditions and Figure 6(ii) represents sets of those objects which does not satisfy the related parametric condition. In Figure 6(ii) we see that no parameter has degree zero thus there is no parameter removal and in Figure 6(ii) no object has zero degree thus there is no requirement to remove any object. Thus, no further reduction is possible.

3.6 Case study and comparative analysis

In this case study, the HOD and faculty of Statistics, MDU University, Rohtak, India wants to add any software in his curriculum for the students. Suggested software's are {SPSS, C++, R, Matlab, C-language, TORA} considering the attributes {Job efficient (JE), Latest(L), Useful in Statistics(US), Easy to Learn (EL), Curriculum Related (CR)}. Based on experts views a soft-set information table is constructed (shown in Table 3). To reduce irrelevant information and assisting in to take the right decision, the proposed algorithm along with some existing algorithm (Maji *et al.*, 2002; Rose *et al.*, 2011) were applied to the case study. Results based on aforesaid algorithm are given in Table 3. According to proposed algorithm it is concluded that the HOD and faculty of the Department may choose SPSS and R as a curriculum. Results obtained by using the algorithm by Maji *et al.* (2002) the choice values, the reduct-soft-set can be represented in Table 3. Here max ci = c1 or c3. Thus, HOD can choose either SPSS or R, whereas based on Rose *et al.* (2011) no column presented as zero significance i.e. no parameter and zero significance. Thus, algorithm has not eliminated or deleted any parameter thus there is no reduction. On Comparing these algorithms, the proposed algorithm and Maji *et al.* (2002) gave same results, but the proposed algorithm eliminates parameters as well as objects in the given information. However the Maji *et al.* (2002) removes only parameters, not objects. On the other hand, Rose *et al.* (2011) is not able to reduce dimension. Thus, it can be concluded that the proposed algorithm is better than algorithms given by both Maji *et al.* (2002) and Rose *et al.* (2011).

4. Conclusions

This paper discusses the problems of dimensionality reduction using soft sets theory and bipartite graphs. An alternative definition of a soft set is discussed, and new algorithms of dimensionality reduction are presented by using proposed techniques which are based on soft sets theory and bipartite graphs. The proposed algorithm eliminates avoidable parameters and objects via parameter and object importance degree. Here we can say that our proposed technique reduces more dimensionality and is easily applicable than the existing ones. Hence, these algorithms execute more proficiently. The proposed algorithm has also been applied to a case study and obtained results were compared with algorithms by Maji *et al.* (2002) and Rose *et al.* (2011) and found that the proposed algorithm is more efficient than existing algorithms. But in various real-life situations the data are not available in binary format of 0 and 1. To overcome this problem, future scope of work involves dimension reduction methods involving hybridization of soft set with other theories.

References

Adams, S. M. (2002). *Biological indicators of aquatic ecosystem stress.* Bethesda, MD: American Fisheries Society.

Atanassov, K. (1986). Intuitionistic fuzzy sets. *Fuzzy Sets and Systems, 20*, 87-96.

Atanassov, K. T. (1994). Operators over interval valued intuitionistic fuzzy sets. *Fuzzy sets and Systems, 64*(2), 159-174.

Alcantud, J. C. R. (2016). Some formal relationships among soft sets, fuzzy sets and their extensions. *International Journal of Approximate Reasoning, 68*, 45–53.

Buriánek, T., Zaorálek, L., Snášel, V., & Peterek, T. (2015). Graph drawing using dimension reduction methods. In A. Abraham, P. Krömer, & V. Snasel (Eds.), *Afro-European Conference for Industrial Advancement. Advances in Intelligent Systems and Computing, Volume 334*. Cham, Switzerland: Springer,

Chen, D., Tsang, E. C. C., Yeung, D. S., & Wang, X. (2005). The parameterization reduction of soft sets and its applications. *Computers and Mathematics with Applications, 49*(5-6), 757-763.

Gupta, P., & Sharma, O. (2015). Feature selection: An overview. *International Journal of Information Engineering and Technology, 1*, 1- 12.

Herawan, T., Rose, A.N.M., & Mat Deris, M. (2009) Soft set theoretic approach for dimensionality reduction. In D. Ślęzak, T. Kim, Y. Zhang, J. Ma, & K. Chung (Eds.), *Database theory and application, DTA 2009, Communications in computer and information science, Volume 64*. Berlin, Heidelberg, Germany: Springer.

Hussain, Z. U., Ali, M. I., & Khan, M. Y. (2016). Application of soft sets to determine Hamilton cycles in a graph. *General Mathematics Notes, 32*. Retrieved from www.i-csrs.orgavailablefreeonlineathttp://www.geman.in.

Table 3. Information table and results from different algorithms

Attributes software	Information table of Software and attributes					Reduced information table using proposed algorithm		Reduced information table using Maji *et al.* (2002)					Reduced information table using Rose *et al.* (2001)				
	JE	L	US	EL	CR	JE	EL	JE	L	US	EL	Choice value	JE	L	US	EL	CR
SPSS	1	1	1	1	1	1	1	1	1	1	1	C1=4	-	-	-	-	-
C++	0	0	1	0	1	-	-	0	0	1	0	C2=1	0	0	1	0	1
R	1	1	1	1	1	1	1	1	1	1	1	C3=4	-	-	-	-	-
Matlab	1	1	1	0	1	1	0	1	1	1	0	C4=3	1	1	1	0	1
C-Language	0	0	1	0	1	-	-	0	0	1	0	C5=1	0	0	1	0	1
TORA	0	1	1	1	1	0	1	0	1	1	1	C6=3	0	1	1	1	1

Jatram, A. & Biswas, B. (2015). Dimension reduction using spectral methods in FANNY for fuzzy clustering of graphs. *2015 Eighth International Conference on Contemporary Computing (IC3), Noida*, 93-96. doi:10.1109/IC3.2015.7346659.

Kamada, T., & Kawai, S. (1989). An algorithm for drawing general undirected graphs. *Information Processing Letters, 31*(1), 7–15.

Lin, T. Y. (1998). Granular computing on binary relations II: Rough set representations and belief functions. *Rough Sets in Knowledge Discovery*, *1*, 122-140.

Ma, X., Zhan, J., Ali, M. I. & Mehmood, N. (2018). A survey of decision making methods based on two classes of hybrid soft set models. *Artificial Intelligence Review, 49*, 511–529. doi:10.1007/s10462-016-9534-2

Maji, P. K., Roy, A. R., & Biswas, R. (2002). An application of soft sets in a decision-making problem. *Computers and Mathematics with Applications, 44*(8-9), 1077-1083.

Mohinta, S., & Samanta, T. K. (2015). An introduction to fuzzy soft graph. *Mathematica Moravica*, *19*(2), 35-48.

Molodtsov, D. (1999). Soft set theory-first results. *Computers and Mathematics with Applications*, *37*, 19-31.

Pawlak, Z. (1982). Rough sets. *International Journal of Information and Computer Science, 11*, 341-356.

Qian, Y., Yin, X., Kong, J., Wang, J., & Gao, W. (2019). Low-rank graph optimization for multi-view dimensionality reduction. *PLoS One 14*(12), e0225987. doi: 10.1371/journal. pone.0225987

Rose, A. N. M., Awang, M. I., Hassan, H., Zakaria, A. H., Herawan, T., & Deris, M. M. (2011) Hybrid reduction in soft set decision making. In D.S. Huang, Y. Gan, V. Bevilacqua, & J. C. Figueroa (Eds.), *Advanced Intelligent Computing, ICIC 2011, Lecture Notes in Computer Science, Volume 6838*. Berlin, Heidelberg, Germany: Springer.

Rosenfeld, A. (1975). Fuzzy graphs. In L. A. Zadeh, K. S. Fu, & M. Shimura (Eds.), *Fuzzy sets and their applications* (pp. 77–95). Cambridge, MA: Academic Press.

Smarandache, F (2018). Extension of soft set to hypersoft set, and then to plithogenic hypersoft set. *Neutrosophic Sets and Systems*, *22*, 168-170.

Thumbakara, R. K., & George, B. (2014). Soft graphs. *General Mathematics Notes*, *21*(2), 75-86.

Yao, Y. Y. (1998). Relational interpretations of neighborhood operators and rough set approximation operators. *Information Sciences*, *111*(1-4), 239-259.

Zadeh, L. A. (1965). Fuzzy sets. *Information and Control, 8*(3), 338-353.

Zhan, J., & Alcantud, J. C. R. (2019). A survey of parameter reduction of soft sets and corresponding algorithms. *Artificial Intelligence Review*, *52*(2019), 1839–1872.

Zhan, J. & Wang, Q. (2019). Certain types of soft coverings based rough sets with applications. *International Journal of Machine Learning and Cybernetics, 10*, 1065-1076.

Zhang, L., Zhan, J. & Alcantud, J. C. R. (2019). Novel classes of fuzzy soft β-coverings-based fuzzy rough sets with applications to multi-criteria fuzzy group decision making. *Soft Computing, 23*, 5327–5351.

Zhang, L. & Zhan, J. (2019). Fuzzy soft β-covering based fuzzy rough sets and corresponding decision-making applications. *International Journal of Machine Learning and Cybernetics, 10*(2019) 1487–1502.

Zhang, L., Zhan, J. M., & Yao, Y. Y. (2020). Intuitionistic fuzzy TOPSIS method based on CVPIFRS models: An application to biomedical problems. *Information Sciences, 517*, 315-339.

Zhao Y., Luo F., Wong S. K. M. & Yao Y. (2007). A general definition of an attribute reduct. In J. Yao, P. Lingras, W. Z. Wu, M. Szczuka, N. J. Cercone, & D. Ślęzak (Eds.), *Rough sets and knowledge technology, RSKT 2007, Lecture notes in computer science, Volume 4481*. Berlin, Heidelberg, Germany: Springer.

Zhi, K., Gao, L., Wang, L., & Li, S. (2008). The normal parameter reduction of soft sets and its algorithm. *Computers and Mathematics with Applications, 56*, 3029–3037.