



# A new parameter reduction algorithm for soft sets based on chi-square test

Hongwu Qin<sup>1</sup> · Qinghua Fei<sup>1</sup> · Xiuqin Ma<sup>1</sup> · Wanghu Chen<sup>1</sup>

Accepted: 5 February 2021 / Published online: 20 March 2021

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC part of Springer Nature 2021

## Abstract

Redundancy is an extremely significant challenge in data integration and decision making based on the model of soft set which is able to process data under uncertainty. There are four available methods which are designed to reduce the redundant parameters of the soft set. But there is a very low success rate on a large number of data sets obtained from practices by these methods. In order to overcome the inherent weakness, we propose a parameter reduction method based on chi square distribution for the model of soft set. Experimental results on two real-life application cases and thirty randomly generated data sets demonstrate that our algorithm largely improves the success rate of parameter reduction, redundant degree of parameter, and has higher practicability in comparison with the four existing normal parameter reduction algorithms.

**Keywords** Soft set · Chi-square test · Redundancy · Parameter reduction

## 1 Introduction

Uncertainty is pervasive in many realistic fields such as economics, engineering, environmental governance, social sciences, business management, etc. Facing these uncertainties, researchers have introduced many mathematical tools, including probability theory, fuzzy set, rough set, etc. However, these mathematical tools have their own shortcomings [1]. Until 1999, Molodtsov, D. [1] put forward a novel soft set theory, which compensates for disadvantages of the insufficient parameterization about classic mathematical methods. It is known to all that soft set is a new powerful mathematical tool for uncertainty.

At present, There are two main aspects in the study of soft set such as theoretical and application research. These articles such as [2–4] discuss some related definitions, properties and

operations of soft sets. There has been a growing interest in addressing the extended models based on soft set. We can list some popular combination models as diverse as fuzzy soft set [5, 6], interval-valued fuzzy soft set [7, 8, 27], integration of the 2-tuple linguistic representation and soft set [9], belief interval-valued soft set [10], confidence soft sets [11], linguistic value soft set [12], separable fuzzy soft sets [13], dual hesitant fuzzy soft sets [14], Z-soft fuzzy rough set [15], Fault-tolerant enhanced bijective soft set [16], and soft rough set [28].

Fortunately, more and more researchers have the resulting interests on developing the applications of soft sets in many areas. Demonstrations of soft set in real applications are easily displayed in the following fields such as information system data analysis [17], decision processing [18–20], conflict processing [21, 22], resource discovery [23], text classification, data mining [24–26] and medical diagnosis so on. In this paper, we address an associated parameter reduction problem for decision making under soft set. Redundancy is an important consideration in data, and parameter reduction is essential for the application of soft set in decision making. Many pioneer researchers have done a lot for this issue. It was firstly suggested in [29] that parameter reduction of soft set was necessary and the related concept was described. Chen et al. [30] improved the parameter reduction of soft sets in [29].

---

✉ Xiuqin Ma  
maxiuqin@nwnu.edu.cn

Hongwu Qin  
qinhongwu@nwnu.edu.cn

<sup>1</sup> College of Computer Science and Engineering, Northwest Normal University, Lanzhou 730070, Gansu, China

**Table 1** Tabular representation of a soft set in example 2.1

$U$	$e_1$	$e_2$	$e_3$	$e_4$	$e_5$
$h_1$	1	0	1	0	1
$h_2$	0	0	0	1	1
$h_3$	0	0	1	1	0
$h_4$	1	1	1	1	1
$h_5$	1	1	1	1	0
$h_6$	1	0	1	0	1

**Table 3** Tabular representations of the soft set  $(F, E)$

$U$	$e_1$	$e_2$	$e_3$	$e_4$	$e_5$
$p_1$	1	1	0	0	1
$p_2$	0	1	0	1	0
$p_3$	1	0	0	0	1
$p_4$	1	1	0	0	1
$p_5$	0	1	1	0	0

However, the method in [30] has the weakness of ignoring the added parameters. So Kong et al. initiated [31] the normal parameter reduction idea without neglecting the newly added parameters. To decrease the computational complexity of this algorithm, the improved methods were created in [32, 33]. However, three above methods for the normal parameter reduction are not appropriate for the large data sets. So a normal parameter reduction algorithm by particle swarm optimization which is suitable for large data sets was expressed in [34, 35]. Next, facing up to the inaccuracy of the method in paper [34], Han et al. [36] again proposed four 0–1 linear programming models. In this paper, we find out that due to its restrictions on special conditions about normal parameter reduction, probability of finding the reduction results and redundant level of parameter on a large number of data sets is very low by means of the above algorithms in [31, 33, 34, 36]. Consequently, we propose a parameter reduction algorithm based on chi square distribution for the soft set, which greatly improves the success rate, redundant degree of parameter and practicality of parameter reduction.

The rest of this paper is structured as follows. Section 2 reviews the basic concepts of soft set theory and four existing methods. Section 3 creates a new parameter reduction method based on chi-square distribution. In Section 4, the proposed chi-square distribution parameter reduction algorithm is applied in

real life, and further compared with the four existing algorithms. Finally, the research conclusions are given in Section 5.

## 2 Basic concepts and the existing normal parameter reduction methods

In this section, we will briefly review the basic concept of soft set theory, which is described through an example and the existing normal parameter reduction methods.

**Definition 2.1** (See [1]) Assume that  $U$  is a non-empty initial universe of objects,  $E$  is a set of parameters in relation to objects in  $U$ , and  $A$  be a subset of  $E$ ,  $\xi(U)$  is the power set of  $U$ . A pair  $(F, A)$  is called a soft set over  $U$ , where  $F$  is a mapping given by

$$F : A \rightarrow \xi(U) \tag{1}$$

In the above definition, the soft set over  $U$  is a parameterized family of subsets of the universe  $U$ . Let’s walk through an example on what is the soft set.

**Example 2.1** An elderly person has a plan to reserve a comfortable nursing house which provides the wonderful old-age service facilities. This elderly person takes into consideration

**Table 2** 2\*2 fourfold table of each pair parameter  $e_\alpha$  and  $e_\beta$

	$e_\beta, f(e_\beta)=0,$	$e_\beta, f(e_\beta)=1$	Sum
$e_\alpha, f(e_\alpha)=0$	$N f(e_\alpha)=0 f(e_\beta)=0$	$N f(e_\alpha)=0 f(e_\beta)=1$	$T_1 = N f(e_\alpha)=0 f(e_\beta)=0$ $+N f(e_\alpha)=0 f(e_\beta)=1$
$e_\alpha, f(e_\alpha)=1$	$N f(e_\alpha)=1 f(e_\beta)=0$	$N f(e_\alpha)=1 f(e_\beta)=1$	$T_2 = N f(e_\alpha)=1 f(e_\beta)=0$ $+N f(e_\alpha)=1 f(e_\beta)=1$
Sum	$T_3 = N f(e_\alpha)=0 f(e_\beta)=0$ $+N f(e_\alpha)=1 f(e_\beta)=0$	$T_4 = N f(e_\alpha)=0 f(e_\beta)=1$ $+N f(e_\alpha)=1 f(e_\beta)=1$	T



**Table 4** Fourfold table of  $e_1, e_2$

	$e_2, f(e_2)=0$	$e_2, f(e_2)=1$	Sum
$e_1, f(e_1)=0$	$0(\frac{2}{5})$	$2(\frac{8}{5})$	2
$e_1, f(e_1)=1$	$1(\frac{3}{5})$	$2(\frac{12}{5})$	3
Sum	1	4	5

similar. We only choose one parameter as representation while the other parameter is redundant. Based on the reduced parameter sets, the decision ability is not changed.

### 3 Parameter reduction method based on chi square distribution for soft set

The parameter reduction method of chi square distribution based on soft set mainly detects redundant parameters by analyzing the correlation between two parameters. In order to obtain the correlation value between the two parameters, we must first calculate the practical correlation frequency and the expected correlation frequency. Next, we will discuss the whole calculation process of the chi square distribution parameter reduction algorithm.

Let the soft set  $(F, E)$  including a set of the object set  $U = \{p_1, p_2, \dots, p_n\}$  and the parameter set  $E = \{e_1, e_2, \dots, e_m\}$ ,  $e_\alpha$  and  $e_\beta$  represent any two parameters in the parameter set  $E$ , and  $f(e_i)(f(e_i) \in 0 \vee 1)$  indicates that the entity value of the soft set. To calculate the practical correlation frequency between parameters  $e_\alpha$  and  $e_\beta$ , we first construct the correlation tables for each pair of parameter  $e_\alpha$  and  $e_\beta$ , that is, fourfold table as shown in Table 2.

In this table, the parameter  $e_\alpha$  and  $e_\beta$  that satisfy the number of entity values  $f(e_i) = 0$  and  $f(e_i) = 1$  are used to form a row (r) or a column (c), the number of parameters  $e_\alpha$  and  $e_\beta$  satisfying the entity values  $f(e_i) = 0$  and  $f(e_i) = 1$  respectively is constructed into a row (r) or column (c), which is represented as  $N_{f(e_i)}$ , that is, the practical correlation frequency between each pair of parameters. For example, the number of parameters that meet the entity value  $f(e_\alpha) = 0$  and  $f(e_\beta) = 0$  between

**Table 5** Fourfold table of  $e_1, e_3$

	$e_3, f(e_3)=0$	$e_3, f(e_3)=1$	Sum
$e_1, f(e_1)=0$	$1(\frac{8}{5})$	$1(\frac{2}{5})$	2
$e_1, f(e_1)=1$	$3(\frac{12}{5})$	$0(\frac{3}{5})$	3
Sum	4	1	5

the parameters  $e_1$  and  $e_2$  is 2 in the dataset of Example 2.1, hence the practical correlation frequency  $N_{f(e_\alpha) = 0, f(e_\beta) = 0}$  between them is 2. The fourfold between  $e_1$  and  $e_2$  is shown in Table 3.

Considering the subsequent work, we use the formula (2) to calculate the degrees of freedom  $V$  in the fourfold table.

$$V = (r-1) * (c-1) \tag{2}$$

From the above fourfold table, it can be seen that the table is a data table with two rows and two columns, and then the degree of freedom  $V = 1$ . According to the calculation results of the above fourfold table, a definition of the chi-square distribution for the correlation between the two parameters is given in detail below.

**Definition 3.1** The correlation between two parameters The correlation between parameters  $e_\alpha$  and  $e_\beta$  is defined as:

$$\begin{aligned} \chi^2_{e_\alpha, e_\beta} &= \sum_{\alpha=i}^c \sum_{\beta=j}^r \frac{(N_{f(e_i)} - T_{f(e_i)})^2}{T_{f(e_i)}} \\ &= \frac{\left( \begin{matrix} N_{f(e_\alpha)=0} - T_{f(e_\alpha)=0} \\ f(e_\beta)=0 & f(e_\beta)=0 \end{matrix} \right)^2}{\begin{matrix} T_{f(e_\alpha)=0} \\ f(e_\beta)=0 \end{matrix}} + \frac{\left( \begin{matrix} N_{f(e_\alpha)=0} - T_{f(e_\alpha)=0} \\ f(e_\beta)=1 & f(e_\beta)=1 \end{matrix} \right)^2}{\begin{matrix} T_{f(e_\alpha)=0} \\ f(e_\beta)=1 \end{matrix}} \\ &+ \frac{\left( \begin{matrix} N_{f(e_\alpha)=1} - T_{f(e_\alpha)=1} \\ f(e_\beta)=0 & f(e_\beta)=0 \end{matrix} \right)^2}{\begin{matrix} T_{f(e_\alpha)=1} \\ f(e_\beta)=0 \end{matrix}} + \frac{\left( \begin{matrix} N_{f(e_\alpha)=1} - T_{f(e_\alpha)=1} \\ f(e_\beta)=1 & f(e_\beta)=1 \end{matrix} \right)^2}{\begin{matrix} T_{f(e_\alpha)=1} \\ f(e_\beta)=1 \end{matrix}} \end{aligned} \tag{3}$$

According to the above-mentioned fourfold table, we can know that  $N_{f(e_i)}$  is the practical correlation frequency between the parameter  $e_\alpha$  and  $e_\beta$ ;  $T_{f(e_i)}$  is the expected correlation fre-

**Table 6** Fourfold table of  $e_1, e_4$

	$e_4, f(e_4)=0$	$e_4, f(e_4)=1$	Sum
$e_1, f(e_1)=0$	$1(\frac{8}{5})$	$1(\frac{2}{5})$	2
$e_1, f(e_1)=1$	$3(\frac{12}{5})$	$0(\frac{3}{5})$	3
Sum	4	1	5

**Table 7** Fourfold table of  $e_1, e_5$

	$e_5, f(e_5)=0$	$e_5, f(e_5)=1$	Sum
$e_1, f(e_1)=0$	$2(\frac{4}{5})$	$0(\frac{0}{5})$	2
$e_1, f(e_1)=1$	$0(\frac{0}{5})$	$3(\frac{3}{5})$	3
Sum	2	3	5

**Table 9** Fourfold table of  $e_2, e_4$

	$e_4, f(e_4)=0$	$e_4, f(e_4)=1$	Sum
$e_2, f(e_2)=0$	$1(\frac{4}{5})$	$0(\frac{1}{5})$	1
$e_2, f(e_2)=1$	$3(\frac{16}{5})$	$1(\frac{4}{5})$	4
Sum	4	1	5

quency between the parameters  $e_\alpha$  and  $e_\beta$ , in other words,  $T_{f(e_i)}$  is the theoretical correlation value between each pair of parameter under the entity values entity values  $f(e_i)=0$  and  $f(e_i)=1$ , which can be calculated by the following formula:

$$T_{f(e_i)} = T_{f(e_{\alpha,\beta})} = \frac{T' \times T''}{T} \tag{4}$$

Combined with the fourfold table,  $T' = \{T_1, T_2\}$  and  $T'' = \{T_3, T_4\}$  respectively indicate the total value of row and column in the case of entity values  $f(e_i)=0$  and  $f(e_i)=1$ . T is

the number of the objects in the soft set. For example 2.1, there are 6 object sets, the T value is 6.

Here, we still use example 2.1, to illustrate the calculation process of formula (3) and (4) in practical applications. Suppose we calculate the correlation between  $e_1$  and  $e_2$  again, we must calculate the the practical correlation frequency value  $N_{f(e_i)}$  and expected correlation frequency value  $T_{f(e_i)}$  a of each pair of parameters before calculating the correlation value  $X^2_{e_{\alpha,\beta}}$  between the two parameters. Since the actual value of each pair of parameters has been calculated in Table 3. Next, we can calculate the expected value of each pair of parameters, as shown below:

$$T_{f(e_{1,2})} = T_{f(e_1)=0, f(e_2)=0} = \frac{T_1 \times T_3}{T} = \frac{\begin{pmatrix} N_{f(e_1)=1} + N_{f(e_1)=0} \\ N_{f(e_2)=0} \quad N_{f(e_2)=1} \end{pmatrix} \times \begin{pmatrix} N_{f(e_1)=1} + N_{f(e_1)=1} \\ N_{f(e_2)=0} \quad N_{f(e_2)=0} \end{pmatrix}}{T} = \frac{2 \times 4}{6} = \frac{8}{6} = \frac{4}{3}$$

$$T_{f(e_{1,2})} = T_{f(e_1)=0, f(e_2)=1} = \frac{T_1 \times T_4}{T} = \frac{\begin{pmatrix} N_{f(e_1)=1} + N_{f(e_1)=0} \\ N_{f(e_2)=0} \quad N_{f(e_2)=1} \end{pmatrix} \times \begin{pmatrix} N_{f(e_1)=1} + N_{f(e_1)=1} \\ N_{f(e_2)=0} \quad N_{f(e_2)=1} \end{pmatrix}}{T} = \frac{2 \times 2}{6} = \frac{4}{6} = \frac{2}{3}$$

$$T_{f(e_{1,2})} = T_{f(e_1)=1, f(e_2)=0} = \frac{T_2 \times T_3}{T} = \frac{\begin{pmatrix} N_{f(e_1)=1} + N_{f(e_1)=1} \\ N_{f(e_2)=0} \quad N_{f(e_2)=1} \end{pmatrix} \times \begin{pmatrix} N_{f(e_1)=0} + N_{f(e_1)=1} \\ N_{f(e_2)=0} \quad N_{f(e_2)=0} \end{pmatrix}}{T} = \frac{4 \times 4}{6} = \frac{16}{6} = \frac{8}{3}$$

$$T_{f(e_{1,2})} = T_{f(e_1)=1, f(e_2)=1} = \frac{T_2 \times T_4}{T} = \frac{\begin{pmatrix} N_{f(e_1)=1} + N_{f(e_1)=1} \\ N_{f(e_2)=0} \quad N_{f(e_2)=1} \end{pmatrix} \times \begin{pmatrix} N_{f(e_1)=0} + N_{f(e_1)=1} \\ N_{f(e_2)=0} \quad N_{f(e_2)=1} \end{pmatrix}}{T} = \frac{4 \times 2}{6} = \frac{8}{6} = \frac{4}{3}$$

**Table 8** Fourfold table of  $e_2, e_3$

	$e_3, f(e_3)=0$	$e_3, f(e_3)=1$	Sum
$e_2, f(e_2)=0$	$1(\frac{4}{5})$	$0(\frac{1}{5})$	1
$e_2, f(e_2)=1$	$3(\frac{16}{5})$	$1(\frac{4}{5})$	4
Sum	4	1	5

**Table 10** Fourfold table of  $e_2, e_5$

	$e_5, f(e_5)=0$	$e_5, f(e_5)=1$	Sum
$e_2, f(e_2)=0$	$0(\frac{2}{5})$	$1(\frac{3}{5})$	1
$e_2, f(e_2)=1$	$2(\frac{8}{5})$	$2(\frac{12}{5})$	4
Sum	2	3	5

**Table 11** Fourfold table of  $e_3, e_4$

	$e_4, f(e_4)=0$	$e_4, f(e_4)=1$	Sum
$e_3, f(e_3)=0$	$3(\frac{16}{5})$	$1(\frac{4}{5})$	4
$e_3, f(e_3)=1$	$1(\frac{4}{5})$	$0(\frac{1}{5})$	1
Sum	4	1	5

Finally, combining with the two parameters of  $e_1$  and  $e_2$  in Table 3, we get the correlation between the two parameters as follows:

$$\begin{aligned}
 X_{e_1, e_2}^2 &= \frac{\sum_{\alpha=1}^c \sum_{\beta=1}^r \frac{(N_{f(e_1)} - T_{f(e_1)})^2}{T_{f(e_1)}}}{\left( \frac{N_{f(e_1)=0} - T_{f(e_1)=0}}{T_{f(e_1)=0}} \right)^2 + \left( \frac{N_{f(e_1)=1} - T_{f(e_1)=1}}{T_{f(e_1)=1}} \right)^2} \\
 &= \frac{\left( \frac{N_{f(e_1)=0} - T_{f(e_1)=0}}{T_{f(e_1)=0}} \right)^2}{\left( \frac{N_{f(e_1)=0} - T_{f(e_1)=0}}{T_{f(e_1)=0}} \right)^2 + \left( \frac{N_{f(e_1)=1} - T_{f(e_1)=1}}{T_{f(e_1)=1}} \right)^2} \\
 &= \frac{\left( \frac{2 - \frac{4}{3}}{\frac{4}{3}} \right)^2}{\left( \frac{2 - \frac{4}{3}}{\frac{4}{3}} \right)^2 + \left( \frac{0 - \frac{2}{3}}{\frac{2}{3}} \right)^2} = \frac{\frac{4}{3}}{\frac{4}{3} + \frac{4}{3}} = \frac{9}{6}
 \end{aligned}$$

**Definition 3.2 correlation matrix** The correlation matrix stores a suite of correlation of all N parameters in Fig. 1, and the correlation between each pair of parameters is represented by  $X_{e_{\alpha}, e_{\beta}}^2$ .

According to the correlation calculation formula between the above two parameters, the correlation between each pair of parameters in example 2.1 is shown in Fig. 2.

For the sake of simplicity, we name our proposed parameter reduction algorithm of chi square distribution as C-SDPR of which detailed steps are given as follows.

**Table 12** Fourfold table of  $e_3, e_5$

	$e_5, f(e_5)=0$	$e_5, f(e_5)=1$	Sum
$e_3, f(e_3)=0$	$1(\frac{8}{5})$	$3(\frac{12}{5})$	4
$e_3, f(e_3)=1$	$1(\frac{2}{5})$	$0(\frac{3}{5})$	1
Sum	2	3	5

**Table 13** Fourfold table of  $e_4, e_5$

	$e_5, f(e_5)=0$	$e_5, f(e_5)=1$	Sum
$e_4, f(e_4)=0$	$1(\frac{8}{5})$	$3(\frac{12}{5})$	4
$e_4, f(e_4)=1$	$1(\frac{2}{5})$	$0(\frac{3}{5})$	1
Sum	2	3	5

**3.1 Our Algorithm: parameter reduction algorithm based on chi square distribution for soft set (C-SDPR)**

- Step 1: Input  $U = \{p_1, p_2, \dots, p_n\}$  and  $E = \{e_1, e_2, \dots, e_m\}$ ;
- Step 2: Calculate the correlation matrix between all two parameters, and check the correlation between parameters.
- Step 3: Access to the rejection hypothesis value D corresponding to the confidence level under the degree of freedom  $V = 1$ . Suppose the correlation between the parameters  $e_{\alpha}$  and  $e_{\beta}$  is higher than D, which are strong correlation between the two parameters. So one of the parameter  $e_{\alpha}$  and  $e_{\beta}$  can be reduced.
- Step 4: Get the new soft set  $(F, E)$  after parameter reduction.

In our algorithm, the confidence level depends on the decision maker itself. The rejection hypothesis value D for judging the correlation between the two parameters is taken from the critical value table of the chi-square distribution, which can usually be found in any statistics textbook. Assuming that the correlation between the two parameters is high, it is sufficient to indicate that the similarity between the two parameters is very high. Hence we can keep one parameter and reduce the other one. In order to interpret our algorithm, we will describe it in detail through an example below.

**Example 3.1** Assume the object set  $U = \{p_1, p_2, p_3, p_4, p_5\}$  and parameter set  $E = \{e_1, e_2, e_3, e_4, e_5\}$ , the mapping relationship of the soft set  $(F, E)$  is shown in Table 3.

According to our method, the following process is shown.

- Step 1: Input the data in the instance.
- Step 2: calculate the fourfold table of each pair of parameters, as shown Tables 4, 5, 6, 7, 8, 9, 10, 11, 12 and 13. Combine the practical correlation frequency value between each pair of parameters in

the fourfold table with formula (4) to calculate the corresponding expected frequency value, which is indicated in brackets of each table. Finally, the correlation between each pair of parameters is calculated by formula (3), and represented in matrix format, as shown in Figs. 3 and 4 respectively.

Step 3: when the degree of freedom  $v = 1$ , we take the confidence level as 0.05 and the corresponding rejection hypothesis D is 3.843 by consulting the chi-square distribution table. From Fig. 3 above mentioned, we can notice that the correlation value of the parameter  $e_1$  and  $e_5$  is  $X_{e_1,5}^2 = 5 > 3.843$ . Accordingly, the two attributes are strong correlation. That is,  $e_1$  or  $e_5$  can be reduced.

Step 4: In the end, we will get the new soft set  $(F, E)$  after parameter reduction as  $\{e_1, e_2, e_3, e_4\}$  or  $\{e_2, e_3, e_4, e_5\}$ .

## 4 The comparison result

In this section, we compare the proposed algorithm with the normal parameter reduction algorithms in [31, 33, 34, 36] on two practical cases in our life and thirty randomly generated datasets, and finally show the comparison results of the five algorithms in tabular form. In short, we term the normal parameter reduction algorithm in [31] as NPR, the algorithm in [33] as NENPR, the normal parameter reduction in soft set based on particle swarm optimization algo-

$$\begin{aligned}
 X_{e_{1,2}}^2 &= \sum_{\alpha=1}^c \sum_{\beta=2}^r \frac{(N_{f(e_{1,2})} - T_{f(e_{1,2})})^2}{T_{f(e_{1,2})}} = \frac{\left\{ \begin{matrix} N_{f(e_1)=0} - T_{f(e_1)=0} \\ f(e_2)=0 \end{matrix} \right\}}{T_{f(e_1)=0} / f(e_2)=0}} + \frac{\left\{ \begin{matrix} N_{f(e_1)=0} - T_{f(e_1)=0} \\ f(e_2)=1 \end{matrix} \right\}}{T_{f(e_1)=0} / f(e_2)=1}} + \frac{\left\{ \begin{matrix} N_{f(e_1)=1} - T_{f(e_1)=1} \\ f(e_2)=0 \end{matrix} \right\}}{T_{f(e_1)=1} / f(e_2)=0}} + \frac{\left\{ \begin{matrix} N_{f(e_1)=1} - T_{f(e_1)=1} \\ f(e_2)=1 \end{matrix} \right\}}{T_{f(e_1)=1} / f(e_2)=1}} \\
 &= \frac{\left(2 - \frac{12}{5}\right)^2}{\frac{12}{5}} + \frac{\left(1 - \frac{3}{5}\right)^2}{\frac{3}{5}} + \frac{\left(2 - \frac{8}{5}\right)^2}{\frac{8}{5}} + \frac{\left(0 - \frac{2}{5}\right)^2}{\frac{2}{5}} = \frac{5}{6} = 0.83 \\
 X_{e_{1,3}}^2 &= \sum_{\alpha=1}^c \sum_{\beta=3}^r \frac{(N_{f(e_{1,3})} - T_{f(e_{1,3})})^2}{T_{f(e_{1,3})}} = \frac{\left\{ \begin{matrix} N_{f(e_1)=0} - T_{f(e_1)=0} \\ f(e_3)=0 \end{matrix} \right\}}{T_{f(e_1)=0} / f(e_3)=0}} + \frac{\left\{ \begin{matrix} N_{f(e_1)=0} - T_{f(e_1)=0} \\ f(e_3)=1 \end{matrix} \right\}}{T_{f(e_1)=0} / f(e_3)=1}} + \frac{\left\{ \begin{matrix} N_{f(e_1)=1} - T_{f(e_1)=1} \\ f(e_3)=0 \end{matrix} \right\}}{T_{f(e_1)=1} / f(e_3)=0}} + \frac{\left\{ \begin{matrix} N_{f(e_1)=1} - T_{f(e_1)=1} \\ f(e_3)=1 \end{matrix} \right\}}{T_{f(e_1)=1} / f(e_3)=1}} \\
 &= \frac{\left(0 - \frac{3}{5}\right)^2}{\frac{3}{5}} + \frac{\left(3 - \frac{12}{5}\right)^2}{\frac{12}{5}} + \frac{\left(1 - \frac{8}{5}\right)^2}{\frac{8}{5}} + \frac{\left(1 - \frac{2}{5}\right)^2}{\frac{2}{5}} = \frac{15}{8} = 1.875 \\
 X_{e_{1,4}}^2 &= \sum_{\alpha=1}^c \sum_{\beta=4}^r \frac{(N_{f(e_{1,4})} - T_{f(e_{1,4})})^2}{T_{f(e_{1,4})}} \\
 &= \frac{\left\{ \begin{matrix} N_{f(e_1)=0} - T_{f(e_1)=0} \\ f(e_4)=0 \end{matrix} \right\}}{T_{f(e_1)=0} / f(e_4)=0}} + \frac{\left\{ \begin{matrix} N_{f(e_1)=0} - T_{f(e_1)=0} \\ f(e_4)=1 \end{matrix} \right\}}{T_{f(e_1)=0} / f(e_4)=1}} + \frac{\left\{ \begin{matrix} N_{f(e_1)=1} - T_{f(e_1)=1} \\ f(e_4)=0 \end{matrix} \right\}}{T_{f(e_1)=1} / f(e_4)=0}} + \frac{\left\{ \begin{matrix} N_{f(e_1)=1} - T_{f(e_1)=1} \\ f(e_4)=1 \end{matrix} \right\}}{T_{f(e_1)=1} / f(e_4)=1}} \\
 &= \frac{\left(0 - \frac{3}{5}\right)^2}{\frac{3}{5}} + \frac{\left(3 - \frac{12}{5}\right)^2}{\frac{12}{5}} + \frac{\left(1 - \frac{8}{5}\right)^2}{\frac{8}{5}} + \frac{\left(1 - \frac{2}{5}\right)^2}{\frac{2}{5}} = \frac{15}{8} = 1.875 \\
 X_{e_{1,5}}^2 &= \sum_{\alpha=1}^c \sum_{\beta=5}^r \frac{(N_{f(e_{1,5})} - T_{f(e_{1,5})})^2}{T_{f(e_{1,5})}} \\
 &= \frac{\left\{ \begin{matrix} N_{f(e_1)=0} - T_{f(e_1)=0} \\ f(e_5)=0 \end{matrix} \right\}}{T_{f(e_1)=0} / f(e_5)=0}} + \frac{\left\{ \begin{matrix} N_{f(e_1)=0} - T_{f(e_1)=0} \\ f(e_5)=1 \end{matrix} \right\}}{T_{f(e_1)=0} / f(e_5)=1}} + \frac{\left\{ \begin{matrix} N_{f(e_1)=1} - T_{f(e_1)=1} \\ f(e_5)=0 \end{matrix} \right\}}{T_{f(e_1)=1} / f(e_5)=0}} + \frac{\left\{ \begin{matrix} N_{f(e_1)=1} - T_{f(e_1)=1} \\ f(e_5)=1 \end{matrix} \right\}}{T_{f(e_1)=1} / f(e_5)=1}}
 \end{aligned}$$

Fig. 3 The correlation between each pair of parameters

$$\begin{aligned}
 &= \frac{\left(3-\frac{9}{5}\right)^2}{\frac{9}{5}} + \frac{\left(0-\frac{6}{5}\right)^2}{\frac{6}{5}} + \frac{\left(2-\frac{4}{5}\right)^2}{\frac{4}{5}} + \frac{\left(0-\frac{6}{5}\right)^2}{\frac{6}{5}} = 5 \\
 X_{e_{2,3}}^2 &= \sum_{\alpha=2}^c \sum_{\beta=3}^r \frac{\left(N_{f(e_{2,3})} - T_{f(e_{2,3})}\right)^2}{T_{f(e_{2,3})}} \\
 &= \frac{\left\{ \frac{N_{f(e_2)=0} - T_{f(e_2)=0}}{f(e_2)=0} \right\}}{T_{f(e_2)=0}} + \frac{\left\{ \frac{N_{f(e_2)=0} - T_{f(e_2)=0}}{f(e_2)=1} \right\}}{T_{f(e_2)=1}} + \frac{\left\{ \frac{N_{f(e_2)=1} - T_{f(e_2)=1}}{f(e_2)=0} \right\}}{T_{f(e_2)=0}} + \frac{\left\{ \frac{N_{f(e_2)=1} - T_{f(e_2)=1}}{f(e_2)=1} \right\}}{T_{f(e_2)=1}} \\
 &= \frac{\left(1-\frac{4}{5}\right)^2}{\frac{4}{5}} + \frac{\left(3-\frac{16}{5}\right)^2}{\frac{16}{5}} + \frac{\left(1-\frac{4}{5}\right)^2}{\frac{4}{5}} + \frac{\left(0-\frac{1}{5}\right)^2}{\frac{1}{5}} = \frac{5}{16} = 0.3125 \\
 X_{e_{2,4}}^2 &= \sum_{\alpha=2}^c \sum_{\beta=4}^r \frac{\left(N_{f(e_{2,4})} - T_{f(e_{2,4})}\right)^2}{T_{f(e_{2,4})}} \\
 &= \frac{\left\{ \frac{N_{f(e_2)=0} - T_{f(e_2)=0}}{f(e_2)=0} \right\}}{T_{f(e_2)=0}} + \frac{\left\{ \frac{N_{f(e_2)=0} - T_{f(e_2)=0}}{f(e_2)=1} \right\}}{T_{f(e_2)=1}} + \frac{\left\{ \frac{N_{f(e_2)=1} - T_{f(e_2)=1}}{f(e_2)=0} \right\}}{T_{f(e_2)=0}} + \frac{\left\{ \frac{N_{f(e_2)=1} - T_{f(e_2)=1}}{f(e_2)=1} \right\}}{T_{f(e_2)=1}} \\
 &= \frac{\left(1-\frac{4}{5}\right)^2}{\frac{4}{5}} + \frac{\left(3-\frac{16}{5}\right)^2}{\frac{16}{5}} + \frac{\left(1-\frac{4}{5}\right)^2}{\frac{4}{5}} + \frac{\left(0-\frac{1}{5}\right)^2}{\frac{1}{5}} = \frac{5}{16} = 0.3125 \\
 X_{e_{2,5}}^2 &= \sum_{\alpha=2}^c \sum_{\beta=5}^r \frac{\left(N_{f(e_{2,5})} - T_{f(e_{2,5})}\right)^2}{T_{f(e_{2,5})}} \\
 &= \frac{\left\{ \frac{N_{f(e_2)=0} - T_{f(e_2)=0}}{f(e_2)=0} \right\}}{T_{f(e_2)=0}} + \frac{\left\{ \frac{N_{f(e_2)=0} - T_{f(e_2)=0}}{f(e_2)=1} \right\}}{T_{f(e_2)=1}} + \frac{\left\{ \frac{N_{f(e_2)=1} - T_{f(e_2)=1}}{f(e_2)=0} \right\}}{T_{f(e_2)=0}} + \frac{\left\{ \frac{N_{f(e_2)=1} - T_{f(e_2)=1}}{f(e_2)=1} \right\}}{T_{f(e_2)=1}} \\
 &= \frac{\left(2-\frac{12}{5}\right)^2}{\frac{12}{5}} + \frac{\left(2-\frac{8}{5}\right)^2}{\frac{8}{5}} + \frac{\left(0-\frac{2}{5}\right)^2}{\frac{2}{5}} + \frac{\left(1-\frac{3}{5}\right)^2}{\frac{3}{5}} = \frac{5}{6} \approx 0.83 \\
 X_{e_{3,4}}^2 &= \sum_{\alpha=3}^c \sum_{\beta=4}^r \frac{\left(N_{f(e_{3,4})} - T_{f(e_{3,4})}\right)^2}{T_{f(e_{3,4})}} \\
 &= \frac{\left\{ \frac{N_{f(e_3)=0} - T_{f(e_3)=0}}{f(e_3)=0} \right\}}{T_{f(e_3)=0}} + \frac{\left\{ \frac{N_{f(e_3)=0} - T_{f(e_3)=0}}{f(e_3)=1} \right\}}{T_{f(e_3)=1}} + \frac{\left\{ \frac{N_{f(e_3)=1} - T_{f(e_3)=1}}{f(e_3)=0} \right\}}{T_{f(e_3)=0}} + \frac{\left\{ \frac{N_{f(e_3)=1} - T_{f(e_3)=1}}{f(e_3)=1} \right\}}{T_{f(e_3)=1}} \\
 &= \frac{\left(0-\frac{1}{5}\right)^2}{\frac{1}{5}} + \frac{\left(1-\frac{4}{5}\right)^2}{\frac{4}{5}} + \frac{\left(3-\frac{16}{5}\right)^2}{\frac{16}{5}} + \frac{\left(1-\frac{4}{5}\right)^2}{\frac{4}{5}} = \frac{5}{16} = 0.3125 \\
 X_{e_{3,5}}^2 &= \sum_{\alpha=3}^c \sum_{\beta=5}^r \frac{\left(N_{f(e_{3,5})} - T_{f(e_{3,5})}\right)^2}{T_{f(e_{3,5})}} \\
 &= \frac{\left\{ \frac{N_{f(e_3)=0} - T_{f(e_3)=0}}{f(e_3)=0} \right\}}{T_{f(e_3)=0}} + \frac{\left\{ \frac{N_{f(e_3)=0} - T_{f(e_3)=0}}{f(e_3)=1} \right\}}{T_{f(e_3)=1}} + \frac{\left\{ \frac{N_{f(e_3)=1} - T_{f(e_3)=1}}{f(e_3)=0} \right\}}{T_{f(e_3)=0}} + \frac{\left\{ \frac{N_{f(e_3)=1} - T_{f(e_3)=1}}{f(e_3)=1} \right\}}{T_{f(e_3)=1}} \\
 &= \frac{\left(0-\frac{3}{5}\right)^2}{\frac{3}{5}} + \frac{\left(1-\frac{2}{5}\right)^2}{\frac{2}{5}} + \frac{\left(1-\frac{8}{5}\right)^2}{\frac{8}{5}} + \frac{\left(3-\frac{12}{5}\right)^2}{\frac{16}{5}} = \frac{15}{8} = 1.875 \\
 X_{e_{4,5}}^2 &= \sum_{\alpha=4}^c \sum_{\beta=5}^r \frac{\left(N_{f(e_{4,5})} - T_{f(e_{4,5})}\right)^2}{T_{f(e_{4,5})}} \\
 &= \frac{\left\{ \frac{N_{f(e_4)=0} - T_{f(e_4)=0}}{f(e_4)=0} \right\}}{T_{f(e_4)=0}} + \frac{\left\{ \frac{N_{f(e_4)=0} - T_{f(e_4)=0}}{f(e_4)=1} \right\}}{T_{f(e_4)=1}} + \frac{\left\{ \frac{N_{f(e_4)=1} - T_{f(e_4)=1}}{f(e_4)=0} \right\}}{T_{f(e_4)=0}} + \frac{\left\{ \frac{N_{f(e_4)=1} - T_{f(e_4)=1}}{f(e_4)=1} \right\}}{T_{f(e_4)=1}} \\
 &= \frac{\left(0-\frac{3}{5}\right)^2}{\frac{3}{5}} + \frac{\left(1-\frac{2}{5}\right)^2}{\frac{2}{5}} + \frac{\left(1-\frac{8}{5}\right)^2}{\frac{8}{5}} + \frac{\left(3-\frac{12}{5}\right)^2}{\frac{16}{5}} = \frac{15}{8} = 1.875
 \end{aligned}$$

Fig. 3 (continued)



**Fig. 4** Correlation matrix for Example 3.1

$$\begin{bmatrix} 0 & & & & & & \\ X^2_{e_2,e_1} & 0 & & & & & \\ X^2_{e_3,e_1} & X^2_{e_3,e_2} & 0 & & & & \\ X^2_{e_4,e_1} & X^2_{e_4,e_2} & X^2_{e_4,e_3} & 0 & & & \\ X^2_{e_5,e_1} & X^2_{e_5,e_2} & X^2_{e_5,e_3} & X^2_{e_5,e_4} & 0 & & \end{bmatrix} = \begin{bmatrix} 0 & & & & & & \\ 0.83 & 0 & & & & & \\ 1.875 & 0.3125 & 0 & & & & \\ 1.875 & 0.3125 & 0.3125 & 0 & & & \\ 5 & 0.83 & 1.875 & 1.875 & 0 & & \end{bmatrix}$$

algorithm in [34] as NPR-PSO, the method in [36] as NPR-LP, respectively. In detail, we make comparison among five methods from two aspects: redundant degree of parameter and success rate.

**A. Case 1: Weather index dataset in different areas**

The weather station wants to predict the weather quality of some areas according to the weather index of different areas in the same period. We collect the meteorological data of some areas from China weather network and transform these data into the style of soft set, which is shown in Table 14. Among them, U represents five different areas in the same period and  $U = \{p_1, p_2, p_3, p_4, p_5\} = \{\text{Lhasa, Dali, Dunhuang, Qingdao, Luoyang}\}$ . E is a set of parameter sets considered and  $E = \{e_1, e_2, e_3, e_4, e_5, e_6\} = \{\text{air pollution index, tourism index, sports index, UV index, temperature index, weather humidity index}\}$ . Next, we will discuss the results of the five algorithms for the above datasets.

(1) Parameter reduction by our algorithm

- Step 1: Input the soft set as shown in Table 14;
- Step 2: According to the given formula and definition, the correlation values between all two parameters are calculated and indicated in matrix form in Fig. 5.
- Step 3: Suppose we take the confidence level is 0.2. When the degree of freedom is 1, the corresponding rejection

hypothesis value D is 1.64 by consulting the chi square distribution table. We find that the correlation between parameter  $e_1$  and  $e_5$ ,  $e_2$  and  $e_3$ ,  $e_2$  and  $e_5$ ,  $e_4$  and  $e_5$  are higher than 1.64, so the parameter  $e_1$  or  $e_5$ ,  $e_2$  or  $e_3$ ,  $e_2$  or  $e_5$ ,  $e_4$  or  $e_5$  can be reduced. After the simplification, the parameters  $\{e_1, e_2, e_4\}$  or  $\{e_3, e_5\}$  can be reduced.

Step 4: Finally, we get the new soft set for minimizing after parameter reduction is  $\{e_3, e_5, e_6\}$ .

(2) Parameter reduction by NPR, NENPR, NPR-PSO and NPR-LP

When the soft set for case 1 is input, we can not discover the parameter subset which satisfies the condition of  $\sum_{e_k \in A} p_{1k} = \sum_{e_k \in A} p_{2k} = \dots = \sum_{e_k \in A} p_{5k}$ . Therefore, we can not find the parameter reduction result by the methods of NPR, NENPR, NPR-PSO and NPR-LP.

(3) Comparison results on case 1.

From the above discussion, the comparison results of the five algorithms are analyzed and presented in Table 15. Here, we give two assessment criteria such as redundant degree of parameter and success rate.

1) Redundant degree of parameter

**Definition 4.1** For soft set (F, E) including a set of the object set  $U = \{p_1, p_2, \dots, p_n\}$  and the parameter set  $E = \{e_1, e_2, \dots, e_m\}$ , the redundant degree of parameter is defined by

$$g = \frac{s}{m} \tag{5}$$

Where  $m$  denotes the number of parameters and  $s$  expresses the number of reduced parameter. Redundant degree of parameter  $g$  represents the ratio of the number of reduced parameter to all of parameter. Notice that the

**Table 14** Soft set for case 1

U	$e_1$	$e_2$	$e_3$	$e_4$	$e_5$	$e_6$
$p_1$	1	0	1	0	1	1
$p_2$	1	1	0	1	0	1
$p_3$	0	0	1	1	1	1
$p_4$	1	1	1	1	0	1
$p_5$	1	0	1	1	0	0





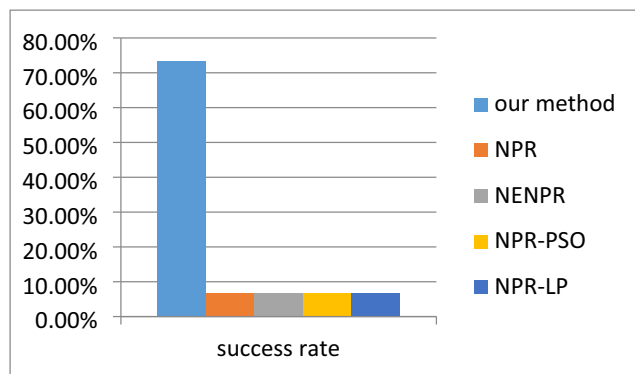
**Table 17** Comparison results on case 2

Algorithm	Parameter reduction	Redundant degree of parameter	Success rate %
Our method	{ $e_3, e_5, e_6$ }	50%	100%
NPR	no result	0	0
NENPR	no result	0	0
NPR-PSO	no result	0	0
NPR-LP	no result	0	0

method on thirty datasets is up to 90.7%; success rate is improved to 90.9%.

### 5 Conclusion

In this paper, we propose a new parameter reduction method based on chi square distribution for the model of soft set. The motivation of this idea is to improve the success rate of finding reduction for the existing methods such as NPR, NENPR, NPR-PSO and NPR-LP. Because of the very low success rate and redundant degree of parameter, the four existing methods are not practical in the real-life applications. On two real cases, success rate of our method is up to 100%, the redundant degree of parameter is up to 50%, in contrast, success rate and redundant degree of parameter of the four existing methods are 0, respectively. The improvement of average redundant degree of parameter of our method on thirty random generated datasets is up to 90.7%; success rate is improved to 90.9% compared with the four existing methods. Finally, we can draw the conclusion that the proposed method has much higher success rate, redun-



**Fig. 7** Success rate of five methods

**Table 18** Comparison results on thirty datasets

Algorithm	Average redundant degree of parameter	Improvement of average redundant degree of parameter	success rate %	Improvement of success rate
Our method	33.2%	90.7%	73.3%	90.9%
NPR	3.1%	90.7%	6.7%	90.9%
NENPR	3.1%	90.7%	6.7%	90.9%
NPR-PSO	3.1%	90.7%	6.7%	90.9%
NPR-LP	3.1%	90.7%	6.7%	90.9%

dant degree of parameter and then practicability compared the existing four approaches.

**Acknowledgements** This work was supported by the National Science Foundation of China (No. 61662067, 61662068, 61762081).

### References

- Molodtsov D (1999) Soft set theory\_First results. *Comput Math Appl* 37(4/5):19–31
- Maji PK, Biswas R, Roy AR (2003) Soft set theory. *Comput Math Appl* 45:555–562
- Ali MI, Feng F, Liu X, Min WK, Shabira M (2009) On some new operations in soft set theory. *Comput Math Appl* 57(9):1547–1553
- Sezgin A, Atagün AO (2011) On operations of soft sets. *Comput Math Appl* 61(5):1457–1467
- Maji PK, Biswas R, Roy AR (2001) Fuzzy soft sets. *Journal of Fuzzy Mathematics* 9(3):589–602
- Ma X, qin H (2018) A distance-based parameter reduction algorithm of fuzzy soft sets. *IEEE Access* 6:10530–10539
- Yang X, Lin TY, Yang J, Dongjun YLA (2009) Combination of interval-valued fuzzy set and soft set. *Comput Math Appl* 58(3): 521–527
- Ma X, Qin H, Sulaiman N, Herawan T, Abawajy JH (2014) The parameter reduction of the interval-valued fuzzy soft sets and its related algorithms. *IEEE Trans Fuzzy Syst* 22(1):57–71
- Wen T, Chang K, Lai H (2020) Integrating the 2-tuple linguistic representation and soft set to solve supplier selection problems with incomplete information. *Eng Appl Artif Intell* 87:103248
- Vijayabalaji S, Ramesh A (2019) Belief interval-valued soft set. *Expert Syst Appl* 119:262–271
- Aggarwal M (2019) Confidence soft sets and applications in supplier selection. *Comput Ind Eng* 127:614–624
- Sun B, Ma W, Li X (2017) Linguistic value soft set-based approach to multiple criteria group decision-making. *Appl Soft Comput* 58: 285–296
- Alcantud JCR, Mathew TJ (2017) Separable fuzzy soft sets and decision making with positive and negative attributes. *Appl Soft Comput* 59:586–595
- Arora R, HarishGarg (2018) A robust correlation coefficient measure of dual hesitant fuzzy soft sets and their application in decision making. *Eng Appl Artif Intell* 72:80–92

15. Zhan J, IrfanAli M, Mehmood N (2017) On a novel uncertain soft set model: Z-soft fuzzy rough set model and corresponding decision making methods. *Appl Soft Comput* 56:446–457
16. Gong K, Wang P, Peng Y (2017) Fault-tolerant enhanced bijective soft set with applications. *Appl Soft Comput* 54:431–439
17. Qin HW, Ma XQ, Herawan T, Zain JM (2012) Dfis: a novel data filling approach for an incomplete soft set. *Int J Appl Math Comput Sci* 22:817–828
18. José C. R. Alcantud, Gustavo S.G. (2017) A new criterion for soft set based decision making problems under incomplete information. *International Journal of Computational Intelligence Systems* 10(1): 394–404
19. Inthumathi V, Chitra V, Jayasree S (2017) The role of operators on soft sets in decision making problems. *International Journal of Computational and Applied Mathematics* 12(3):899–910
20. Maji PK, Roy AR (2002) An application of soft sets in a decision making problem. *Comput Math Appl* 44(8-9):1077–1083
21. Sutoyo E, Mungad M, Hamid S, Herawan T (2016) An efficient soft set-based approach for conflict analysis. *PLOS ONE*
22. Haruna K, Ismail MA, Suyanto M, Gabralla LA et al (2019) A soft set approach for handling conflict situation on movie selection. *IEEE Access* 7:116179–116194
23. Ezugwu AE, Adewumi AO (2017) Soft sets based symbiotic organisms search algorithm for resource discovery in cloud computing environment. *Futur Gener Comput Syst* 76:33–55
24. Herawan T, Mat Deris M (2011) A soft set approach for association rules mining. *Knowl-Based Syst* 24(1):186–195
25. Feng F, Cho J, Pedrycz W, Fujita H, Herawan T (2016) Soft set based association rule mining. *Knowl-Based Syst* 111:268–282
26. Qin H, Ma X, Zain JM, Herawan T (2012) A novel soft set approach for selecting clustering attribute. *Knowl-Based Syst* 36:139–145
27. Ma X, Qin H (2020) A new parameter reduction algorithm for interval-valued fuzzy soft sets based on Pearson's product moment coefficient. *Appl Intell* 50:3718–3730
28. Zhan J, Liu Q, Herawan T (2017) A novel soft rough set: soft rough hemirings and corresponding multicriteria group decision making. *Appl Soft Comput* 54:392–402
29. Maji PK, Biswas R, Roy AR (2002) An application of soft sets in a decision making problem. *Comput Math Appl* 44(8–9):1077–1083
30. Chen D, Tsang ECC, Yeung DS, Wang X (2005) The parameterization reduction of soft sets and its applications. *Comput Math Appl* 49(5–6):757–763
31. Kong Z, Gao L, Wang L, Li S (2008) The normal parameter reduction of soft sets and its algorithm. *Comput Math Appl* 56(12):3029–3037
32. Xie N (2016) An algorithm on the parameter reduction of soft sets. *Fuzzy Information and Engineering* 8(2):127–145
33. Ma X, Sulaiman N, Qin H, Herawan T, Zain JM (2011) A new efficient Normal parameter reduction algorithm of soft sets. *Comput Math Appl* 62(2):588–598
34. Kong Z, Jia W, Zhang G, Wang L (2015) Normal parameter reduction in soft set based on particle swarm optimization algorithm. *Appl Math Model* 39(16):4808–4820
35. Han B (2016) Comments on “Normal parameter reduction in soft set based on particle swarm optimization algorithm”. *Appl Math Model* 44(23–24):10828–10834
36. Han B, Li Y, Geng S (2017) 0–1 linear programming methods for optimal normal and pseudo parameter reductions of soft sets. *Appl Soft Comput* 54:467–484

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Hongwu Qin** received PhD in computer science from Universiti Malaysia Pahang in 2012, and M.Sc. in computer science from Beijing University of Technology in 2005. He is currently an associate professor of College of Computer Science and Engineering, Northwest Normal University. He has published more than 20 research papers. His current research interests include big data and soft set.