This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2021.3086030, IEEE Access

Date of publication xxxx 00, 0000, date of current version xxxx 00, 0000.

Digital Object Identifier 10.1109/ACCESS.2017.Doi Number

Environmental efficiency evaluation of the Xiangjiang River basin: A DEA cross-efficiency approach with social network

Zhaokun Cheng [1], Fenfen Li [2], and Xiangyang Tao [1]

[1] School of Business, Central South University, Changsha 410083, China
[2] School of Public Administration, Hunan University of Finance and Economics, Changsha, 410205, China

Corresponding author: Fenfen Li (e-mail: lifenfen@hufe.edu.cn).

The research is supported by The Youth Foundation of Social Science and Humanity of China Ministry of Education (Grant No. 19YJC630074), National Natural Science Foundation of China (No. 71871223), In-novation-Driven Planning Foundation of Central South University (2019CX0410).

ABSTRACT Cross-efficiency in data envelopment analysis is widely used for performance evaluation of decision-making units (DMUs), but it neglects the reference relationship between each pair of DMUs. In order to address the problem, this paper introduces the concept of social network to clearly describe the reference relationship between DMUs and then a fully ranking can be made. A novel cross-efficiency approach is developed by integrating social network analysis in this study. Firstly, we propose a pairwise comparison model based on cross-efficiency evaluation to identify the superiority and inferiority between any pair of DMUs. Secondly, based on pairwise comparison, we build a non-weighted directed social network where a direction generates from one DMU to another DMU if the former one references the latter one. Each edge in the unweighted social network captures the learning procedure from worse-performance DMU to better-performance DMU. Comparing with traditional cross-efficiency approaches, our proposed approach considers the reference relationship among DMUs rather than only cross-efficiency scores. Thus, the importance of each DMU in the network can be measured by its centrality. Finally, the proposed approach is employed to environmental efficiency evaluation of the cities along the Xiangjiang River Basin, then, some beneficial universal policies have been summarized.

INDEX TERMS data envelopment analysis; cross-efficiency; social network; Xiangjiang River Basin

I.INTRODUCTION

Performance evaluation of entities is necessary for decision makers to judge their pros and cons and further find solutions to improve their performance. Data envelopment analysis (DEA) proposed by Charnes et al. [1] is a data-driven approach that can measure the efficiency of decision-making units (DMUs) with multiple inputs and outputs. DEA is widely used in the ranking of banks, universities, and other entities. Sung et. al. [2] used DEA to analyses the operational efficiency of the research aspect of the graduate university, In addition, there are some other studies that use DEA in efficiency evaluation in different fields [3,4]. However, traditional DEA models, i.e., CCR and BCC models [1,5], turned out to be defective in ranking DMUs, as they lacked discrimination of DEA efficient DMUs.

Traditional DEA model is a self-evaluation model since it allows DMUs to select weights independently for evaluation, in which weights may be impractical or unreasonable for other DMUs. By considering the importance of peers in the evaluation process, the cross-efficiency approach was proposed which can eliminate the influence of unrealistic weights in self-evaluation models and enhance the ability to rank DMUs. However, due to the no uniqueness of the optimal weights obtained by cross-efficiency model, the DMU cross-efficiency scores are non-unique, which reduces the accuracy of the evaluation results.

In this study, by combining social network analysis, we develop a novel DEA cross-efficiency approach using pairwise comparison for ranking. To the best of our knowledge, current DEA approaches rarely consider the relationship between DMUs, even though the relationship can affect evaluation efficiencies and improvement policy. Liu et al. [6] constructed a network by applying the λ values in classical DEA models as the strength of the network link, and the final rank result was obtained via the concept of eigenvector centrality in social network analysis. Liu and Lu

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2021.3086030, IEEE Access

[7] extended the ranking approach in Liu et al. [6] by removing the bias caused by a scale difference among organizations, and employed the new approach to rank the R&D performance of Taiwan's government-supported research institutes. de Blas et al. [8] proposed a new ranking approach by combining DEA and the concept of "measures of dominances" in social network analysis, which used a weighted Hyperlink-Induced Topic Search algorithm to describe the authorities and the hubs in social network. Recently, Ang et al. [9] developed a modified approach to select benchmarks combining social network analysis, from which the ranking results were also obtained. More discrimination could be provided when DEA combining to social network analysis [10,11], in the study of Umut et. al. [12], social network-based eigenvector centrality values were used as the weights of the super-efficiency scores. The social network attributes was regarded as an evaluation indicators in some studies. Simone et. al. [13] measured the extent of employment contracts selection effectively converted into influential social network positioning, Margarita et. al. [14] analysed the relationship between productive efficiency and online-social-networks (OSN) considering several indicators of business "social Media" activities.

The proposed approach in this study is different from previous studies from the following three aspects. Firstly, DEA cross-efficiency rather than classical DEA efficiency measures (i.e., Farrell efficiency measures) is applied to build the corresponding social network for ranking. Previous studies employ the intensity vector in DEA (envelopment form) to build social network, while this study uses the weights of inputs and outputs in DEA (multiplier form) to calculate self-evaluation and peer-evaluation scores for constructing social network. Secondly, this study ranks DMUs by judging the importance of DMU relative to the overall performance evaluation network not the individual performance. DEA cross-efficiency scores are used to judge the performance of any pair DMUs, rather than directly used to determine the ranking results. In fact, the ranking result is determined by "centrality". Thirdly, different from the directed and weighted social network in previous studies, the proposed approach ranks DMUs by constructing a directed and unweighted social network.

Note that the proposed ranking approach also helps with the benchmark selection. The direction in the constructed social network represents a learning process for performance improvement, that is, nodes (DMUs) with worse performance benchmark toward nodes with better performance through the identified directions. Nodes receive more directions from other nodes implies that the node is selected as benchmarks for more nodes. In addition, the direction can be regarded as a certain learning process, where a series of directions can be formulated to one benchmarking path for the social network [15,16]. Therefore, compared with traditional DEA-based benchmarking studies such as Ramón et al. [17] and An et al. [18,19], our approach provides a new idea for sequential benchmarking by considering learning relationship between DMUs in social network, and provide more ways to obtain efficiency improvement strategies.

The rest of this paper is organized as follows. Section 2 provides the preliminaries where the CCR model and the DEA cross-efficiency model are reviewed. Section 3 shows the proposed DEA cross-efficiency evaluation model based on social-network. Section 4 presents a numerical example to verify the rationality of the proposed model. The proposed approach is applied to environmental efficiency of Xiangjiang River Basin in Section 5. Finally, conclusions and further research directions are given in Section 6.

II. PRELIMINARIES

Before introducing our new approach, we review the CCR model and cross-efficiency model.

A. CCR MODEL

Assume that there are n DMUs that consume m inputs to generate s outputs. Denote the ith input and rth output for $DMU_j\ (j = 1, 2, \dots, n)$ as $x_{ij}\ (i = 1, 2, \dots, m)$ and $y_{rj}\ (r = 1, 2, \dots, s)$ respectively. The CCR model is described as follows.

where $v_{id}\ (i = 1,2,\dots,m)$ and $u_{rd}\ (r = 1,2,\dots,s)$ are the weights attached to the inputs and outputs of DMU_d respectively. Denote the optimal solution of (1) as $\{v_{id}^*, u_{rd}^*, \forall i, \forall r\}$. $E_{dd} = \frac{\sum_{r=1}^{s} u_{rd}^* y_{rd}}{\sum_{i=1}^{m} v_{id}^* x_{id}}$ is defined as the efficiency of DMU_d. Since each DMU always finds its

$$\begin{aligned}
\max \quad & E_{dd} = \frac{\sum_{r=1}^{s} u_{rd} y_{rd}}{\sum_{i=1}^{m} v_{id}\, x_{id}} \\
s.t. \quad & \sum_{r=1}^{s} u_{rd} y_{rj} - \sum_{i=1}^{m} v_{id} x_{ij} \le 0, \\
& \qquad j = 1,2,\dots,n, \\
& v_{id} \ge 0, \quad i = 1,2,\dots,m, \\
& u_{rd} \ge 0, \quad r = 1,2,\dots,s
\end{aligned} \tag{1}$$

optimal weights for optimizing its efficiency, the efficiency of a DMU by using CCR model is called self-evaluation efficiency. DMU_d is DEA efficient if $E_{dd} = 1$.

Since (1) is a fraction programming that is complicated to solve, it can be transformed into the following linear programming.

$$\begin{aligned}
max \quad & E_{dd} = \sum_{r=1}^{s} \mu_{rd} y_{rd} \\
s.t. \quad & \sum_{r=1}^{s} \mu_{rd} y_{rj} - \sum_{i=1}^{m} \omega_{id} x_{ij} \le 0, \\
& \qquad j = 1,2 \dots, n \\
& \sum_{i=1}^{m} \omega_{id}\, x_{id} = 1, \\
& \omega_{id} \ge 0, \quad i = 1,2,\dots,m, \\
& \mu_{rd} \ge 0, \quad r = 1,2,\dots,s.
\end{aligned} \tag{2}$$

B. CROSS-EFFICIENCY MODEL

Sexton et al. [20] developed a cross-efficiency model based on the aforementioned (2). First, by solving (2), we can obtain the DMU_d 's optimal weights $\{\omega_{id}^*, \dots, \omega_{md}^*, \mu_{rd}^*, \dots, \mu_{sd}^*\}$. Then, the efficiency of DMU_j $(j = 1, 2, \dots, n)$ with respect to DMU_d is calculated by

$$E_{dj} = \frac{\sum_{r=1}^{s} \mu_{rd}^* y_{rj}}{\sum_{i=1}^{m} \omega_{id}^* x_{ij}} \tag{3}$$

Through Formula (3), each DMU_d obtains a self-evaluation efficiency E_{dd} and $n-1$ peer-evaluation efficiencies E_{dj} $(j = 1,2, \dots, j-1, j+1, \dots, n)$. Then, the cross-efficiency of DMU_j $\overline{E_j}$ is defined as the average of self-evaluation efficiency and peer-evaluation efficiencies.

$$\overline{E_j} = \frac{1}{n}\sum_{d=1}^{n} E_{dj}, \qquad j = 1,2, \dots, n. \tag{4}$$

Note that the optimal weights obtained from (1) may not be unique for each DMU_d. To solve this problem, Doyle and Green [21] proposed the benevolent and aggressive cross-efficiency models which can be shown as (5) and (6), respectively.

$$\begin{aligned}
Max \quad & \sum_{r=1}^{s} \mu_{rd}\Big(\sum_{j=1, j\neq d}^{n} y_{rj}\Big) \\
s.t. \quad & \sum_{i=1}^{m} \omega_{id} x_{ij} - \sum_{r=1}^{s} \mu_{rd} y_{rj} \geq 0, \\
& \qquad\qquad j = 1,2 \dots, n \\
& \sum_{i=1}^{m} \omega_{id}\Big(\sum_{j=1, j\neq d}^{n} x_{ij}\Big) = 1, \\
& \sum_{r=1}^{s} \mu_{rd} y_{rd} - E_{dd}\sum_{i=1}^{m} \omega_{id} x_{id} = 0, \\
& \omega_{id} \geq 0, \quad i = 1,2, \dots, m, \\
& \mu_{rd} \geq 0, \quad r = 1,2, \dots, s.
\end{aligned} \tag{5}$$

$$\begin{aligned}
Min \quad & \sum_{r=1}^{s} \mu_{rd}\Big(\sum_{j=1, j\neq d}^{n} y_{rj}\Big) \\
s.t. \quad & \sum_{i=1}^{m} \omega_{id} x_{ij} - \sum_{r=1}^{s} \mu_{rd} y_{rj} \geq 0, \\
& \qquad\qquad j = 1,2 \dots, n \\
& \sum_{i=1}^{m} \omega_{id}\Big(\sum_{j=1, j\neq d}^{n} x_{ij}\Big) = 1, \\
& \sum_{r=1}^{s} \mu_{rd} y_{rd} - E_{dd}\sum_{i=1}^{m} \omega_{id} x_{id} \\
& \qquad = 0, \\
& \omega_{id} \geq 0, \quad i = 1, \dots, m, \\
& \mu_{rd} \geq 0, \quad r = 1, \dots, s.
\end{aligned} \tag{6}$$

In Models (5) and (6), E_{dd} is the CCR efficiency of DMU_d generated from (1). (5) aims to maximize the sum of other DMUs' peer-evaluation efficiency scores, while (6) seeks to minimize the sum of other DMUs' peer-evaluation efficiency scores. Note that both the two models optimize the sum of other DMUs' peer-evaluation efficiency scores keeping the efficiency of DMU_d at its self-evaluation level (CCR efficiency).

III. THE PROPOSED RANKING APPROACH

The proposed ranking approach consists of three steps. Firstly, we propose a new cross-efficiency evaluation model based on pairwise comparison. Secondly, we construct a social network for ranking DMUs with the proposed cross-efficiency evaluation model. Thirdly, we introduce the concept of "centrality measures" in social network analysis to derive the ranking results. The proposed approach is summarized graphically in figure 1. The next subsections illustrate the new approach in detail.

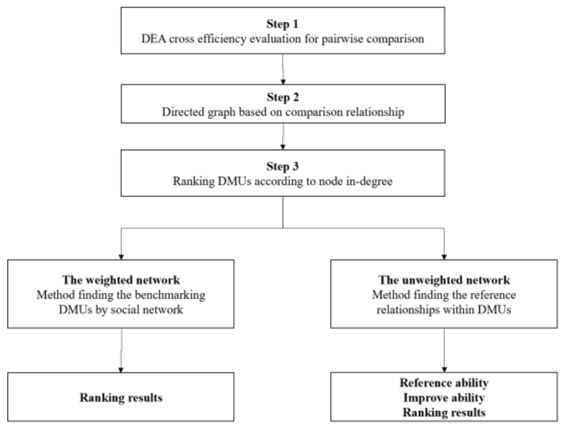

FIGURE 1. The summary of the proposed approach.

A. CROSS-EFFICIENCY EVALUATION MODEL USING PAIRWISE COMPARISON

Although the original cross-efficiency model can rank efficient DMUs, it suffers from non-unique cross-efficiency scores, which reduces the stability of the ranking results. Besides, traditional DEA cross-efficiency model considers the peer-evaluation among DMUs, but it suffers from two shortcomings for constructing a learning relationship as the relationship between each pair of DMUs cannot be well identified independently in social network. In the traditional cross-efficiency evaluation, peer-evaluation between two DMUs actually considers the effect of other DMUs which is contained in the constraints of the programming model. Thus, the traditional DEA cross-efficiency model may be not well-suited for describing the learning relationship. In this paper, we propose a new cross-efficiency model to build the learning relationship between each pair of DMUs in social network.

Consider any pair DMUs in social network, DMU_p and DMU_q $(p = 1,2, \dots, n; q = 1,2, \dots, n; q \neq p)$. Referring to the original DEA cross-efficiency model, DMU_p and DMU_q should make self-evaluation and peer-evaluation when the overall evaluation system only incorporate one pair of DMUs, i.e., DMU_p and DMU_q . Accordingly, the self-evaluation model of DMU p using pairwise comparison is expressed in the following.

$$\begin{aligned} max \quad & E_{pp} = \sum_{r=1}^{s} \mu_{rp} y_{rp} \\ s.t. \quad & \sum_{r=1}^{s} \mu_{rp} y_{rp} - \sum_{i=1}^{m} \omega_{ip} x_{ip} \le 0, \quad p = 1,2 \dots, n \\ & \sum_{r=1}^{s} \mu_{rp} y_{rq} - \sum_{i=1}^{m} \omega_{ip} x_{iq} \le 0, \quad q = 1,2, \dots, n, \; q \ne p \\ & \sum_{i=1}^{m} \omega_{ip} x_{ip} = 1 \\ & \omega_{ip} \ge 0, \quad i = 1,2, \dots, m \\ & \mu_{rp} \ge 0, \quad r = 1,2, \dots, s \end{aligned} \tag{7}$$

Denote the optimal solution of Model (7) as $\{\omega_{ip}^*, \mu_{rp}^*, \forall i \forall r\}$. $E_{pp} = \sum_{r=1}^{s} \mu_{rp}^* y_{rp}$ is the self-evaluation score of DMU_p. Note that Model (7) only incorporates two inequality constraints where the efficiency scores of DMU_p and DMU_q are restricted at interval $(0,1]$. It can be found that the original DEA cross-efficiency model has n inequality constraints while the new model has only two inequality constraints.

Similarly, the self-evaluation model of DMU q using pairwise comparison is formulated in the following.

$$\begin{aligned} max \quad & E_{qq} = \sum_{r=1}^{s} \mu_{rq} y_{rq} \\ s.t. \quad & \sum_{r=1}^{s} \mu_{rq} y_{rq} - \sum_{i=1}^{m} \omega_{iq} x_{iq} \le 0, \quad q = 1,2 \dots, n \\ & \sum_{r=1}^{s} \mu_{rq} y_{rp} - \sum_{i=1}^{m} \omega_{iq} x_{ip} \le 0, \quad p = 1,2, \dots, n, \; p \ne q \\ & \sum_{i=1}^{m} \omega_{iq} x_{iq} = 1 \\ & \omega_{iq} \ge 0, \quad i = 1,2, \dots, m \\ & \mu_{rq} \ge 0, \quad r = 1,2, \dots, s \end{aligned} \tag{8}$$

Denote the optimal solution of (8) as $\{\omega_{iq}^*, \mu_{rq}^*, \forall i, \forall r\}$. $E_{qq} = \sum_{r=1}^{s} \mu_{rq}^* y_{rq}$ is the self-evaluation score of DMU_q. Based on the optimal solution of (7) and (8), the peer-evaluation efficiency scores of DMU_p and DMU_q can be described accordingly as follows.

$$E_{pq} = \frac{\sum_{r=1}^{s} \mu_{rp}^* y_{rq}}{\sum_{i=1}^{m} \omega_{ip}^* x_{iq}} \tag{9}$$

$$E_{qp} = \frac{\sum_{r=1}^{s} \mu_{rq}^* y_{rp}}{\sum_{i=1}^{m} \omega_{iq}^* x_{ip}} \tag{10}$$

where E_{pq} denotes the peer-evaluation efficiency score of DMU_q from DMU_p and E_{qp} denotes the peer-evaluation efficiency score of DMU_p from DMU_q.

B SOCIAL NETWORK WITH LEARNING RELATIONSHIP

Based on the new DEA cross-efficiency evaluation model using pairwise comparison, the learning relationship can be identified and the social network can be constructed. In this paper, all DMUs under ranking are regarded as nodes in social network. The edges represent the learning relationship that DMUs with worse performance will learn from DMUs with better performance. Additionally, given a network, one of the most important problems in social network analysis is the identification of important key nodes. A common approach for ranking nodes in network is to use centrality measures as the centrality measure of a node reflects the node's importance in the whole network [22]. Thus, this paper adopts the concept of centrality to obtain the final ranking results.

As the cross-efficiency scores of any pair DMUs are derived from the DEA cross-efficiency evaluation model using pairwise comparison, the next task is to determine the weight and direction of each edge in the constructed social network. In this paper, we denote the cross-efficiency from pairwise comparison as the weight between the evaluated pair DMUs, and the corresponding direction is pointed from the peer-evaluation DMU to the self-evaluation DMU.

Consider two nodes, note p and note q (i.e., DMU_p and DMU_q). Their self-evaluation efficiency scores and peer-evaluation efficiency scores are obtained through (7), (8) and (9), (10). In social network analysis, the four efficiency scores formulate the association matrix A_{pq}, which is shown as follows.

$$A_{pq} = \begin{bmatrix} E_{pp} & E_{pq} \\ E_{qp} & E_{qq} \end{bmatrix} \tag{11}$$

On the basis of A_{pq}, we can infer the weights of node p and q by the following procedure. First, considering that there are two weights between nodes p and q, defined by w_{pq} and w_{qp}, the four components of the association matrix should be synthetized into two weights. As the cross-efficiency scores can be regarded as comprehensive evaluations that take into account the peer-evaluation and self-evaluation of the pair DMUs, we adopts the cross-efficiency scores as the weights between nodes p and q.

The weight from node q to node p is denoted as the cross-efficiency score of DMU p under pairwise comparison, that is,

$$w_{qp} = E_p = \frac{E_{pp} + E_{qp}}{2} \tag{12}$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2021.3086030, IEEE Access

The weight from node p to node q is denoted as the cross-efficiency score of DMU q under pairwise comparison, that is,

$$w_{pq} = E_q = \frac{E_{pq} + E_{qq}}{2} \tag{13}$$

Denote the edge from node p to node q as $e(p, q)$. Then, the cross-efficiency scores under pairwise comparison can formulate $n(n-1)$ edges in social network. Through formulas (12) and (13), we can describe the weights of all edges $e(p, q), p = 1, \ldots, n; q = 1, \ldots., n, p \neq q$ in the social network.

C SOCIAL NETWORK RANKING APPROACH

1) THE CENTRALITY MEASURES

Based on the social network constructed above, the next task is to rank nodes from the concept of social network analysis. In this paper, nodes are ranked through the network degree center attribute. In social network analysis, the in-degree of node p is the summary of connections from other nodes to node p. It reflects the importance of node p in the whole social network. The in-degree measure of node p is defined as follows.

$$S_p = \alpha \sum_{l}^{n} w_{lp} \tag{14}$$

where w_{lp} denotes weight of edge $e(l, p)$, S_p is the in-degree measure which denotes the centrality degree of node p and α is a constant. The in-degree measure is computed as the sum of the centrality value (weight) of its neighbors multiplied by a constant α.

2) THE WEIGHTED NETWORK

We focus on centrality measures for dominance (or reference) networks where relationships between nodes are weighted and directed through a dominance relation. Since we use the cross-efficiency scores under pairwise comparison as weights for any pair DMUs, the in-degree measure can be reformulated as follows.

$$S_p = \sum_{q}^{n} w_{qp} = \sum_{q}^{n} \bar{E}_{qp} \tag{15}$$

For simplicity, α is set 1 in Formula (15). Since $\bar{E}_{qp}$ is a binary variable, the value of in-degree measure will be an integer. The larger S_p is, the higher performance of node p will be. The final ranking result is derived by ranking S_p for all node $p,\ p = 1,2, \ldots, n$.

This ranking method is to add the weight values of all edges pointing to the node p, and use the sum of the weight values as the sorting basis.

$$R'^{(p)} = S_p \tag{16}$$

3) THE UNWEIGHTED NETWORK

In practice, each DMU can learn from any DMU which has better performance than it. If a DMU has a better performance than another DMU in the new cross-efficiency evaluation, then the latter one has the willingness to learn from the former one. Based on this point, we show a simplification principle.

If $w_{qp} = E_p \geq w_{pq} = E_q$, then let $w'_{qp} = 1$ and $w'_{pq} = 0$. w'_{qp}, w'_{pq} are defined by the new weights of edges $e(q, p)$ and $e(p, q)$ for the social network. w'_{qp} is a binary variable that describes whether node q learns from node p. If $w'_{qp} = 1$, node q will learn from node p; if $w'_{qp} = 0$, node q will not learn from node p. Similar explanation can be applied to $\bar{E}_{pq}$. For ease of illustration in the late sections, we call this transformation by 0-1 transformation.

The principle above simplifies the social network by reassigning weights between any pair nodes at 0 or 1. For any pair nodes in the social network, the node that has higher cross-efficiency score under pairwise comparison will be assigned weight with 1, while the node that has lower cross-efficiency score under pairwise comparison will be assigned weight with 0. Note that edges with weight 0 can be deleted from the social network. Therefore, the simplified social network only incorporates $n(n-1)/2$ edges that have the same weights with 1. In addition, since edges direct to nodes with better performance from nodes with worse performance, the edges reveal the learning relationship that nodes with worse performance can learn from nodes with better performance.

Except of ranking DMUs, the proposed approach can also provide the information on benchmarking, particularly, the benchmarking path selection. Benchmarking is a managerial instrument to improve performance in the business world, in which DMUs' performance should be measured first. To illustrate how the proposed approach can be used for benchmarking, we introduce the following definition based on the concept of social network analysis.

Definition 1. *Referability*. Referability of node p denotes the degree of reference from other nodes, viz. referability of node p equals to the in-degree of node p in social network analysis. Mathematically, referability of node p is denoted as $R(p)$, then we have $R(p) = \sum_q^n w_{qp} = \sum_q^n \bar{E}_{qp}$.

Definition 2. *Improvability*. Improvability of node p denotes the degree of improvement to other nodes, that is, improvability of node p equals the out-degree of node p in social network analysis. Mathematically, improvability of node p is denoted as $I(p)$, then we have

$$I(p) = \sum_{q}^{n} w_{pq} = \sum_{q}^{n} \bar{E}_{pq} \tag{17}$$

Actually, in the context of benchmarking, DMU p with larger $R(p)$ implies that more DMUs select it as benchmark. That is, more DMUs learn from it to improve their performance. DMU p with larger $I(p)$ indicates that it can select more DMUs to learn from in performance improvement. Therefore, for any DMU p, a series of edges can be found directing to other nodes, and these edges formulate the benchmarking path in performance improvement.

In addition, given that the social network is formulated by pairwise comparison, we can observe the following characteristic of referability and improvability.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2021.3086030, IEEE Access

Theorem 1. For any node p in the constructed social network, we have $R(p)+I(p)=n-1$.

Proof. Through the DEA cross-efficiency evaluation model using pairwise comparison, we can infer that node p will be compared with other $n-1$ nodes. Since node p either directs to other nodes or is directed, we can infer that the sum of in-degree measure and out-degree measure is $n-1$. Thus, $R(p)+I(p)=n-1$. ■

A DMU referenced by more DMUs has more universality, and is more suitable as targets for the phased improvement for some DMUs. Besides, in the social network, a DMU is referenced by more DMUs means that it is in a more important position. Thus, the DMU will has more influence in the evaluation.

IV. A numerical example and an application to monitoring stations along the Xiangjiang River Basin

A A NUMERICAL EXAMPLE

We first consider a numerical example with 10 DMUs that use three inputs to generate two outputs. The data is randomized between interval [1,10]. TABLE I reports the data set of 10 DMUs.

TABLE I
THE DATA SET OF 10 DMUS

DMU	x_1	x_2	x_3	y_1	y_2
1	4	9	7	5	6
2	9	4	8	3	7
3	6	4	8	3	7
4	9	4	4	8	5
5	5	5	7	6	8
6	8	3	8	9	6
7	3	3	5	6	9
8	3	7	6	7	6
9	8	5	7	3	5
10	4	3	4	8	7

The ranking results of different approaches are summarized in **TABLE Ⅱ**, where the cross-efficiency ranking approach, the proposed pairwise evaluation approach using weighted network, the transform approach using unweighted network are reported. It can be inferred that the results obtained by the social network ranking are similar as those obtained by the traditional DEA cross-efficiency ranking. Compared with the ranking of performance efficiency, the ranking obtained by our proposed approach after the 0-1 transformation in section C can reflect referential significance. A DMU ranking higher is more referenceable.

TABLE Ⅱ
RANKING RESULT USING DIFFERENT APPROACHES

DMU	RANK (cross-efficiency)	Efficiency (cross-efficiency)	RANK (pairwise evaluation)	Efficiency (pairwise evaluation)	RANK (P)	R (P)	RANK (CCR)	Efficiency (CCR)
1	8	0.4357	8	0.6452	8	2	7	0.5884
2	9	0.4328	9	0.6296	9	1	8	0.5833
3	7	0.4497	7	0.6730	7	3	9	0.5833
4	4	0.6726	6	0.7335	5	5	2	1.0000
5	6	0.5821	3	0.8179	6	5	6	0.6381
6	3	0.6878	4	0.7825	3	6	1	1.0000
7	1	0.9600	2	0.9291	1	9	4	1.0000
8	5	0.6692	5	0.7699	4	6	3	1.0000
9	10	0.3221	10	0.5596	10	0	10	0.3968
10	2	0.9528	1	0.9627	2	8	5	1.0000

When calculating the efficiency of the traditional CCR method, there is a problem of insufficient recognition of efficient DMUs. As shown in TABLE Ⅱ， several DMUs are efficient, and the order of efficient DMUs cannot be distinguished correctly when sorting.

When the proposed method is applied to the comparison of a small number of DMUs, all DMUs can be sorted and DMUs that perform well can be distinguished, but the top DMUs are prone to the problem of similar scores and cannot be further distinguished. In order to rank the DMUs that perform well more clearly, a weight screening method can be introduced here to enlarge the weight difference. When the weight of a continuous edge of a node representing a DMU is less than 1, its weight is changed to 0, and only the edge with a weight of 1 can be retained.

In the pairwise comparison method combining social network and cross-efficiency, proposing the 0-1 transformation or not represents different meanings. When edge weights are included in the calculation of centrality, the ranking of network nodes is more inclined to comparison of the degree of effectiveness. A decision-making unit that repeatedly wins in pairwise comparison means that its efficiency is higher than that of other decision-making units. When edge weights are not included in the calculation of centrality, that is, when the weighted network is converted to the weightless network, the ordering of network nodes desires reference universality.

B An application to monitoring stations along the Xiangjiang River Basin

The Xiangjiang River is a significant tributary of the Yangtze River. It plays a leading role in Hunan Province in economic growth. As for the population, 41.5 million people settled down in the Xiangjiang River Basin, accounting for 60.89% of Hunan Province's population. Moreover, the largest hydropower station with the best regulation performance in Hunan Province, namely, Dongjiang Hydropower Station, is also located in the Xiangjiang River Basin. The region's GDP adds up to 2.4855 trillion Yuan in 2016, accounting for 78.8% of the total GDP of Hunan Province. The current paper collects the data of monitoring stations along Xiangjiang

River and investigates environmental efficiencies and rankings of these stations using the proposed approach.

1) INPUTS, DESIRABLE OUTPUTS AND UNDESIRABLE OUTPUT OF THE XIANGJIANG RIVER BASIN

This paper chooses GDP and river pollution as the economic and environmental indicators, respectively. Input indicators generally contain capital investment, labor input, and energy input. Particularly, energy input is typically represented by standard coal consumption. Except for regular inputs, the current study also encounters undesirable outputs as inputs. The selected outputs are the comprehensive pollution indexes of water quality, including Dissolved Oxygen (DO), Total Phosphorus (TP), Metals, and pH.

Dissolved oxygen is the oxygen content dissolved in water, which is a vital indicator of aerobic aquatic organisms' living conditions. Most aquatic organisms need a specific concentration of oxygen to breathe and metabolize normally, and the living environment above or below this level may have adverse effects on their survival. Emissions of pollutants containing large amounts of organic matter (e.g., waste from untreated paper, food processing, and other industries) can significantly reduce dissolved oxygen concentrations. Even if the low dissolved oxygen condition does not last long, it may lead to aquatic life's death. Within a specific range, the higher the fraction of dissolved oxygen, the better. Therefore, dissolved oxygen is selected as desirable output in current paper.

Phosphorus in water may be present in elemental phosphorus, orthophosphate, condensed phosphate, pyrophosphate, metaphosphate, and organic group-bound phosphate. Its primary sources are domestic sewage, chemical fertilizers, organophosphorus pesticides, and phosphate builders used in modern detergents. The eutrophication of waters is closely related to the nitrogen and phosphorus contents in water bodies. Phosphorus is the main nutrient element that affects the growth and reproduction of algae. Therefore, the total phosphorus is regarded as an undesirable output in this study.

Metals exist in the earth's crust and are released into the environment through physical and chemical weathering of rocks. The geological characteristics of the watershed determine the concentration of metals in water bodies. However, human-made metal sources also exist, including industrial waste, domestic waste, agricultural pollution, and anti-fouling coatings for some ships. Most metals are toxic to organisms above an absolute concentration, and some metals accumulate in animals and plants. These metals can enter the food chain through surface contact, breathing, and feeding. Therefore, Metals are supposed to be an undesirable output in this paper.

pH measures the acidity or alkalinity of water, ranging from acidic (pH<7) to neutral (pH=7) and alkaline (pH>7). Most aquatic organisms require a living environment with a specific pH range. If the water is too alkaline or too acidic, the organism's life activities will be destroyed. Alkaline and acidic substances are put into water bodies in the natural environment, but human activities often cause pH changes in water bodies (e.g., chemical leakage and mine waste discharge). Besides, the acidic water quality will make it easier for organisms to absorb metal poisons, which increases the possibility of bioconcentration and will have severe impacts on the ecosystem. The standard pH range is from 6 to 9; the Xiangjiang River data are all in the range of 7 to 8. The outstanding value of the pH value is 7; the smaller the deviation, the better. In this paper's data, the pH value is greater than 7; thus, PH is supposed to be an undesirable output in the current paper.

In general, the inputs selected in environmental efficiency evaluation are

x_1 Total number of employees (TNE; 1000 persons)

x_2 Total fixed asset investment in industry (TFA; 1billion yuan)

x_3 Total energy consumption in that region (TEC; 10,000 tons of standard coal equivalent

the desirable outputs are

y_1 Gross domestic product in industry (GDP; 1 billion yuan)

y_2 Dissolved oxygen (DO; mg/L)

and undesirable outputs are

z_1 Permanganate index (COD Mn; mg/L)

z_2 Ammonia-nitrogen (AN; mg/L)

z_3 Concentration of Pb (Pb; $\mu g/L$)

z_4 Total phosphorus (TP; mg/L),

z_5 The content arsenic among water (As; $\mu g/L$)

z_6 Cd concentration (Cd; mg/L)

z_7 Pondus hydrogenii (pH)

2) RESULTS AND DISCUSSIONS

The current paper selects 34 monitoring stations along Xiangjiang River Basin as DMUs, and taking water pollution as an undesired output, the environmental efficiency and ranking results are obtained.

Sorting the 34 regions through different analysis methods can obtain information on both the efficiency value and the reference ability.

TABLE III
EFFICIENCIES AND RANKS IN DIFFERENCE APPROACHES

DMU number	Region	RANK (CCR)	Efficiency (CCR)	RANK (cross-efficiency)	Efficiency (cross-efficiency)
1	Dongan	13	1.0000	19	0.9133
2	Lengshuitan	8	1.0000	6	0.9603
3	Shuangpai	20	1.0000	24	0.8471
4	Lingling	23	1.0000	1	0.9898
5	Qidong	3	1.0000	2	0.9684
6	Changning	15	1.0000	9	0.9489
7	Yanfeng	29	0.9661	28	0.8205
8	Hengdong	25	0.9991	18	0.9167

9	Changning	28	0.9759	21	0.9038
10	Hengyang	17	1.0000	16	0.9206
11	Shigu	34	0.7082	34	0.6072
12	Leiyang	32	0.9423	27	0.8313
13	Leiyang 2	31	0.9442	25	0.8352
14	Hengdong 2	18	1.0000	10	0.9488
15	Hengdong 3	10	1.0000	11	0.9480
16	Xintian	26	0.9991	31	0.8065
17	Lukou	16	1.0000	20	0.9101
18	Lukou 2	24	1.0000	17	0.9182
19	Lusong	22	1.0000	22	0.8974
20	Xiangtan	12	1.0000	5	0.9639
21	Xiangtan 2	2	1.0000	8	0.9542
22	Xiacheng	14	1.0000	30	0.8089
23	Yuetang	11	1.0000	29	0.8095
24	Yuhu	33	0.9161	26	0.8322
25	Tianxin	4	1.0000	4	0.9646
26	Yuelu	19	1.0000	3	0.9664
27	Wangcheng	27	0.9778	23	0.8711
28	Liuyang	1	1.0000	32	0.7786
29	Changsha twon	30	0.9469	33	0.6531
30	Guiyang	6	1.0000	7	0.9547
31	Zixing	21	1.0000	12	0.9455
32	Zixing 2	7	1.0000	13	0.9428
33	Zixing 3	5	1.0000	15	0.9328
34	Xiangyin	9	1.0000	14	0.9351

TABLE III
EFFICIENCIES AND RANKS IN DIFFERENCE APPROACHES

DMU number	Region	R(P)'	RANK (transform)	R(P)
1	Dongan	22.1774	2	29
2	Lengshuitan	23.3007	4	28
3	Shuangpai	18.6146	12	21
4	Lingling	24.6164	1	33
5	Qidong	23.4981	7	25
6	Changning	20.0506	20	13
7	Yanfeng	18.8011	25	9
8	Hengdong	19.8740	24	10
9	Changning	18.4030	28	6
10	Hengyang	21.9523	8	25
11	Shigu	17.8931	33	1
12	Leiyang	18.5813	30	4
13	Leiyang 2	19.1536	29	5
14	Hengdong 2	22.6996	11	22
15	Hengdong 3	22.1730	16	20
16	Xintian	18.3930	26	9
17	Lukou	19.8848	23	11
18	Lukou 2	20.2004	21	13
19	Lusong	19.9091	13	21
20	Xiangtan	21.7644	9	23
21	Xiangtan 2	21.0878	14	21
22	Xiacheng	17.8649	31	4
23	Yuetang	17.7909	32	2
24	Yuhu	21.0594	19	17
25	Tianxin	20.8010	15	21
26	Yuelu	20.3196	17	20
27	Wangcheng	19.7135	27	8
28	Liuyang	17.9703	22	13
29	Changsha twon	17.6244	34	1
30	Guiyang	19.8606	10	23
31	Zixing	20.6428	5	28
32	Zixing 2	20.5851	3	29
33	Zixing 3	20.2420	6	26
34	Xiangyin	21.2679	18	20

Through weighted social network analysis, it can be obtained that the efficiency of region 4(Lingling) is better and its rank is higher. In weighted social network analysis, the relationship between DMUs is a pairwise comparison, and the strength of the reference relationship is reflected through the weight. Region 4 has more weighted in-degrees, which indicates that region 4(Lingling) performs better in pairwise cross-comparison with more DMUs, and some DMUs have a larger gap with region 4. That is, the higher weighted in-degree indicates the comprehensive advantages of area 4 both in terms of the number and strength of relationships. Therefore, region4(Lingling) is a typical benchmark under the CCR method.

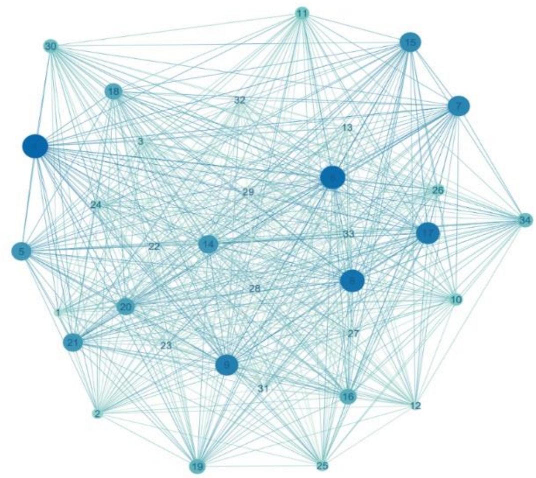

FIGURE 2. Reference relationship of weighted social network

Region 1(Dongan) performed in terms of efficiency value and general reference value, which is a typical inefficient unit. The decision-making unit represented by Region 1(Dongan) has the characteristic of not ranking high in efficiency value, but ranking high in reference, which shows that there are such types of benchmarks in the evaluation, and their performance is relative to that of decision-making units with high efficiency values. It is not prominent, but it can become a reference object for other decision-making units with a slight advantage. Region 4(Lingling) still ranks first in the ranking methods of unweighted networks, which explains the rationality of the unweighted network method to a certain extent, because the region 4(Lingling) has the best performance and can be referenced by almost all DMUs.

However, the weight of edges between nodes only reflects the gap between two regions, not shows the direction of improvement. Because not all DMUs are able to refer to the DMU with a greater gap.

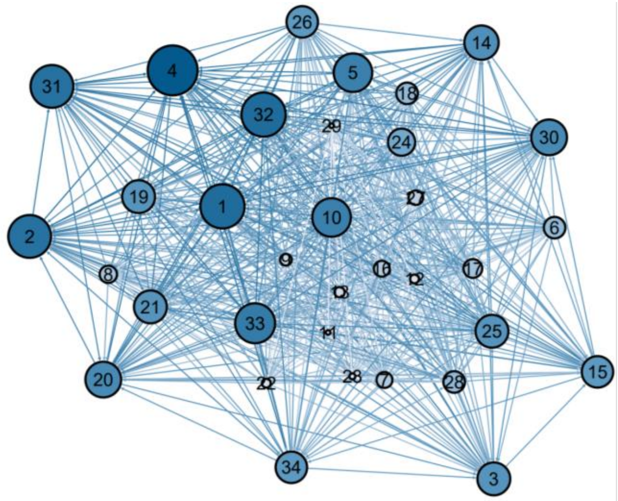

FIGURE 3. **The reference relationship of DMUs without weights**

Combining the value of centrality and the reference relationship of the network diagram, the town can choose its own reference object. The dominant DMU has a point that points to the directed edge of the target DMU. As shown in FIGURE 3., the region 14(Hengdong2) can select any node with edge in unweighted social network as reference object, such as region4(Lingling), region 5(Qidong), region 6(Changning), region 7(Yanfeng), region 8(Hengdong1), region 9(Hengdong1)and region 15(Hengdong3). Lingling has best performance, but Hengdong3 can be the nearest improvement target.

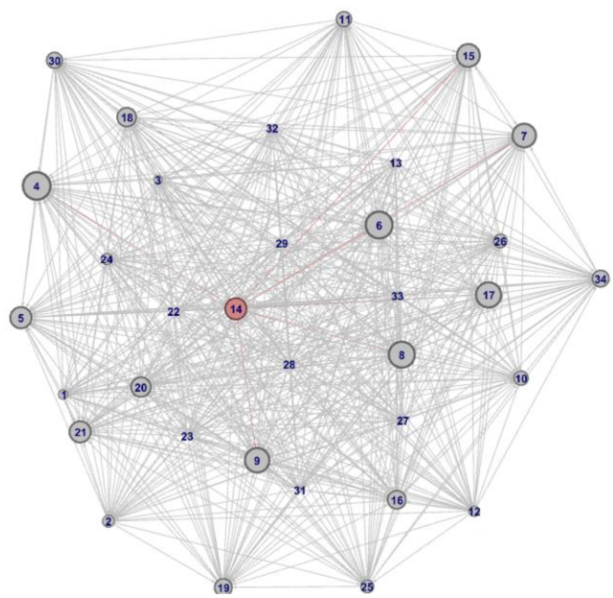

FIGURE 4. **Selection of reference points for specific townships**

When it is necessary to set learning objects for most of DMUs, examples with universal reference significance are more valuable. After analyzing the actual situation of the Xiangjiang River Basin with universal reference significance, the following results can be obtained.

(1) The regions that rank low in CCR but top in the unweighted network, such as Zixing, correspond to the Xiangjiang River basin city Chenzhou, which has a prominent performance in environmental protection policies, especially in the treatment of heavy metal pollution.

(2) In areas where the city's reference ability ranking and the traditional CCR ranking are low, the existing problems can be summarized from the three aspects of heavy metal pollution, agricultural pollution, and domestic pollution. Looking for references from the unauthorized network diagram, you can find some environmental efficiency. For surrounding areas with higher values, select cities with close geographical locations from the reference points and policy improvements are more practical.

(3) In the past, urban development has been accompanied by environmental pollution. In fact, areas with higher levels of economic development tend to have lower environmental efficiency than other areas. Attention should still be paid to the balance between economic growth and environmental protection.

V. POLICY IMPLICATIONS AND CONCLUSIONS

The method proposed in this paper combines cross-efficiency evaluation with social network, and has excellent performance in improving the discrimination of ranking, and provides a method for setting learning objects.

The pollution sources of the Xiangjiang River can be roughly divided into industrial pollution and agricultural pollution. Starting from the environmental protection policies of areas with higher rankings in the above analysis, Lingling, Dongjiang, and Chenzhou, we have obtained the following policy recommendations.

As an inland river, under the classification of industrial pollution, heavy metal pollution is a sharp manifestation of the contradiction between the environmental protection and the economic development. Heavy metal pollution mainly involves non-ferrous, chemical, smelting and other industries, and the pollution sources are mainly concentrated in Qingshuitang of Zhuzhou, Yuetang of Xiangtan, Shuikoushan of Hengyang and Industrial Zone of Zhubugang. To solve heavy metal pollution, the first step is to restrict high-pollution risk industries, the second step is to carry out technological transformation of high-pollution enterprises, the third step is to centrally manage enterprises in high-pollution industries, and the fourth step is to disclose information on enterprise pollution for restraint. Concentrated areas of heavy metal-related enterprises such as Shuikoushan in Hengyang along the Yangtze River closed down polluting enterprises and realized industrial transformation.

In addition, high-tech monitoring should be applied to some watershed ecological safety protection measures. In Changsha, a real-time online water quality inspection system has been installed since 2014, and water quality inspection equipment has been installed at the outlets to grasp sewage information in real time.

Organic pollutants in water bodies may be caused by agricultural non-point source pollution or industrial pollution. In recent years, the pollution of the livestock and poultry

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2021.3086030, IEEE Access

breeding industry has become more and more serious, which has caused the concentration of organic pollution to increase. The first solution is to adjust the industrial structure, eliminate outdated production capacity policies, and change the high proportion of the heavy chemical industry in the region. The second solution is to transform production methods and promote circular transformation in industrial parks, and explore ways to change industrial production methods with high energy consumption and high emissions.

The application of environmental protection policies and high-tech in industry and agriculture is imminent. We need to change the economic development model by transforming the traditional mass production model into a sustainable, green and low-carbon model. Local governments should actively formulate environmental regulations and increase investment in pollution control. low environmental performance areas can select areas with learning relationships in social network, and learn improvement measures from areas that has similar actual situation. In addition, areas that have been referenced a lot but not DEA efficient are likely to have universal reference value in some policies. environment protection measures of them may improve the overall ecological benefits of the Xiangjiang River.

There are also some limitations in this paper. When calculating the efficiency of neighboring cities, the interaction between upstream and downstream rivers is not considered. In actual efficiency evaluation, upstream cities will inevitably affect downstream cities. In addition, this article does not consider some factors that may affect the assessment of environmental efficiency, such as the development of tourism, the cultural level of the labor force, and the degree of agricultural mechanization. More methods are needed to improve this research in the future.

REFERENCES

[1] A. Charnes, W. W. Cooper, and E. Rhodes, Measuring the efficiency of decision making units. *Eur J Oper Res*, vol. 2, no. 6, pp. 429-44, 1978.

[2] Z. Liu, E. Wu, Y. Qiao, et al, Performance evaluation of public bus transportation by using dea models and Shannon? s entropy: An example from a company in a large city of China. *IEEE/CAA Journal of Automatica Sinica.*, 2020.

[3] E. Filiopoulou, P. Mitropoulou, N.Lionis, and C. Michalakelis, On the efficiency of cloud providers: A DEA approach incorporating categorical variables. *IEEE Transactions on Cloud Computing.*,vol. 7, pp. 149, 2017.

[4] L. Chia-Yen, W. Ke, and S. Wenqiang, Allocation of emissions permit for China's iron and steel industry in an imperfectly competitive market: A Nash equilibrium DEA method. *IEEE Transactions on Engineering Management.*, January 2020.

[5] R. D. Banker, A. Charnes, and W. W. Cooper. Some models for estimating technical and scale inefficiencies in data envelopment analysis. *Manage Sci*, vol. 30. no 9, pp.1078-92, 1984.

[6] J. S. Liu, W-M. Lu, C. Yang, et al. A network-based approach for increasing discrimination in data envelopment analysis. *J Oper Res Soc*, vol. 60, no. 11, pp.1502-10, 2009.

[7] J. S. Liu, and W-M, Lu. DEA and ranking with the network-based approach: a case of R&D performance. *Omega-Int J Manage Sci*, vol. 38, no 6, pp.453-64, 2010.

[8] C. S. D. Blas, J. S. Martin, and D G Gonzalez. Combined social networks and data envelopment analysis for ranking. *Eur J Oper Res*, vol. 266, no. 3, pp.990-9, 2018.

[9] S. Ang, R. Zheng, F. Wei, et al, A modified DEA-based approach for selecting preferred benchmarks in social networks. *J. Oper. Res. Soc.* 2019.

[10] S. Lozano, and L. Calzada-Infante, Computing gradient-based stepwise benchmarking paths. *Omega-Int. J. Manage. Sci*. vol. 81, pp. 195–207, 2018.

[11] A. Ghahraman, and D. Prior, A learning ladder toward efficiency: Proposing network-based stepwise benchmark selection. *Omega-Int. J. Manage. Sci*., vol. 63, pp. 83-93, 2016.

[12] U. Aydın, M. A. Karadayi, and F. Ülengin, How efficient airways act as role models and in what dimensions? A superefficiency DEA model enhanced by social network analysis. *Air Transp. Manag.*, vol. 82, pp. 101725, 2020.

[13] D. Oliveira Gobbo, S. C, E. B. Mariano, et al, Combining social network and data envelopment analysis: A proposal for a Selection Employment Contracts Effectiveness index in healthcare network applications. *Omega-Int. J. Manage,* 2020.

[14] M. Martínez-Núñez, and W. S. Pérez-Aguiar, Efficiency analysis of information technology and online social networks management: An integrated DEA-model assessment. *Inf. Manage*., vol. 51, no. 6, pp. 712-725, 2014.

[15] S. Lozano, and L.Calzada-Infante, Efficiency assessment using network analysis tools. *J. Oper. Res. Soc*. vol. 69, no.11, pp.1803–1818, 2018.

[16] S. Lozano, and L. Calzada-Infante, Efficiency ranking using dominance network and multi-objective optimization indexes. *Expert Syst. Appl*., vol. 126, pp. 83-91, 2019.

[17] N. Ramón, Ruiz, J. L, and I. Sirvent, Cross-benchmarking for performance evaluation: Looking across best practices of different peer groups using DEA. *Omega-Int. J. Manage. Sci*., vol. 92, pp. 102169, 2020.

[18] Q. An, X. Tao, and B. Xiong, Benchmarking with data envelopment analysis: an agency perspective. *Omega-Int. J. Manage. Sci*., 2021.

[19] Q. An, X. Tao, B. Dai, et al, Bounded-change target-setting approach: Selection of a realistic benchmarking path. *J. Oper. Res. Soc.*, 2020.

[20] T. R. Sexton, R. H. Silkman, and A. J. Hogan, Data envelopment analysis: Critique and extensions. *New Directions for Program Evaluation*., vol. 32, pp.73-105, 1986.

[21] J. Doyle, and R. Green, Efficiency and cross-efficiency in DEA: Derivations, meanings and uses. *J. Oper. Res. Soc*., vol. 45, no.5, pp.567-578, 1994.

[22] L. C. Freeman, Centrality in social networks conceptual clarification. *Soc. Networks*., vol. 1, no. 3, pp. 215-239, 1978.

FIRST A. Zhaokun Cheng Ms. Zhaokun Cheng is a master student of Management Science and Engineering in the School of Business, Central South University, China. She received the B.S. degree in industrial engineering from China University of mining and technology, Xuzhou, China, in 2018. Her research interests include cross-efficiency, productivity measurement and improvement, social network analysis.

SECOND B. Fenfen Li Fenfen Li is currently an assistant professor in Hunan University of Finance and Economics, China. Her main research interest includes sustainable management, environmental governance, performance evaluation, etc. She has published about 20 papers in journals and conference proceedings on the above topic.

THIRD C. Mr. Xiangyang Tao is a PhD student of Management Science and Engineering in the School of Business, Central South University, China. His research interests include benchmarking, productivity measurement and improvement, the principal-agent model and auction. He has published 5 papers in prestigious journals such as Omega, Journal of Operational Research Society, International Transactions in Operational Research, Computational Economics and Environmental Science and Pollution Research.