Supplementary Material A

A. Spectral Clustering

Spectral clustering is a clustering method based on spectral graph theory, which treats the clustering task as a graph partitioning problem. It constructs an undirected weighted graph with each sample being a vertex and the similarity between any two samples being the weight of the edge connecting the two vertices. Then the graph is divided into several sub-graphs, namely clusters, by certain graph cut method. The final purpose is to find a partition of the graph such that the edges between different sub-graphs have low weights, while edges within the same sub-graph have high weights, indicating that samples in different clusters are different from each other and samples within the same cluster are similar to each other.

There are several traditional graph cut methods, such as minimum cut, ratio cut, normalized cut and min/max cut, etc. According to different graph cut methods, spectral clustering can be divided into two categories: iterative spectral clustering and multi-path spectral clustering. Among them, the multi-path spectral clustering has been more widely used because of its simple and easy-to-understand characteristics, and the NJW algorithm is one of the classic multi-path spectral clustering algorithms.

Let $\{x_1, x_2, \ldots x_n\}$ be a set of samples that needs to be clustered into k clusters, then the process of NJW algorithm is described as follows:

Step 1 Construct an undirected weighted graph according to the relationships (e.g., similarity) among samples, and W is set as its adjacency matrix.

Step 2 Calculate the normalized Laplacian matrix L_{sym}. First, calculate the degree of each vertex as $d_i = \sum_j w_{ij}$, and a diagonal matrix with the degrees of all vertices as the diagonal elements is derived, which is called the degree matrix D. Then, the Laplace matrix can be derived as $L = D - W$, and the normalized Laplacian matrix L_{sym} can also be derived by normalizing L as $L_{sym} = D^{-\frac{1}{2}} L D^{-\frac{1}{2}}$.

Step 3 Calculate the eigenvectors corresponding to the first k largest eigenvalues $v_1, v_2, \ldots, v_k$ of L_{sym}.

Step 4 Arrange $v_1, v_2, \ldots, v_k$ in the column direction to form a new matrix $V \in R^{n \times k}$, and standardize the new matrix V by rows to obtain a new matrix $U \in R^{n \times k}$.

Step 5 Take each row in U as a sample in a k-dimensional new solution space, use K-means to cluster these samples, and map the clustering results back to the original solution space to obtain the final clusters.

To sum up, spectral clustering maps the samples to a lower-dimensional space through the Laplacian matrix. In the low-dimensional data space, the samples have better clustering characteristics (relations among the data points have been well preserved in this new space). Compared with traditional clustering methods such as K-means clustering, spectral clustering does not make any assumptions about the global structure of the data, and has the ability to identify non-convex distributed clusters, which is very suitable for many practical problems. Moreover, since social network is essentially a graph, the spectral clustering designed for partitioning a graph into sub-graphs is well-motivated to be used to identify the structure of the social network.

B. CRP in Multi-attribute LSGDM Problems

Multi-attribute LSGDM refers to a process in which more than 20 DMs evaluate multiple attributes of a set of feasible alternatives and select the best solution from them. Since there are a large number of stakeholders with different backgrounds involved in LSGDM, they have different interest preference and may have divergences on some issues. Therefore, CRP is particularly important in multi-attribute LSGDM, which helps to narrow the gap among opinions of DMs until a consensus is reached and prevents stakeholders from opposing the decision result. A typical consensus model in multi-attribute LSGDM consists of the following four phases:

(1) Preference expression phase. Let $E=\{e_1,e_2,\ldots,e_q\}$ be a set of DMs, who evaluate several attributes $R=\{r_1,r_2,\ldots,r_n\}$ of a set of alternatives $X=\{x_1,x_2,\ldots,x_m\}$ and give their preference matrices $P_k=\{p_{ij}^k\}_{m\times n}$, $k=1,2,\ldots,q$. There are many forms for DMs to express their evaluation information, such as interval numbers, fuzzy sets, intuitionistic fuzzy sets and hesitating fuzzy sets. Among them, IFSs exhibit a sound ability to deal with the uncertainty caused by fuzziness and subjectivity, and could be used in LSGDM problems.

(2) Clustering phase. DMs are divided into several sub-groups according to the similarity between their characteristics [19], such as their personal opinions, social relationship, etc. The complexity of the LSGDM problem is largely reduced by clustering the large-scale DMs into several sub-groups which are easier to be analyzed.

(3) Consensus measurement phase. The consensus degree of the group is measured to judge whether the DMs have reached an agreement. There are two commonly used consensus measures [20], [44], which are either based on the distances to the collective preference or based on the distances between DMs [2].

(4) Preference recommendation-feedback phase. If the consensus degree does not meet the requirements, the feedback mechanism will be activated to help DMs modify their opinions. There are two main types of feedback mechanisms: (a) the one based on identification rules and direction rules [45], which identifies DMs contributing less to the consensus and provides them with the direction to modify their preference; and (b) the one based on optimization rules [23], [33], [34], [46], which aims to minimize the adjustment costs (e.g., the number of adjusted DMs, alternatives, or the distance between the original preference and the adjusted preference).

Based on the aforementioned four phases, this article proposes a novel two-stage consensus model. Before formally introducing the proposed model, we first list the main notations that would be used.

TABLE II MAIN NOTATIONS USED IN THE PROPOSED MODEL

Notation	Explanation
e_k	k-th DM
C_l	l-th cluster
x_i	i-th alternative
r_j	j-th attribute
p_{ij}^k	DM e_k 's preference for attribute r_j of alternative x_i
λ_{kh}	e_k 's initial trust in e_h
s_{kh}	The preference similarity between e_k and e_h
TD_{kh}	e_k 's hybrid trust degree in e_h
VI	The validity index of clustering
β	The consensus threshold
θ	The adjustment parameter for clusters
ρ	The adjustment parameter for individual DMs
G_i	The score of the alternative x_i

C. ***Sensitivity Analysis***

1) Sensitivity Analysis on Consensus Threshold

In the proposed model, the consensus threshold is determined empirically. In order to explore the impact of the consensus threshold on the proposed model, different consensus thresholds will be substituted into the model to observe the performance of some key indicators of the consensus model.

With the same input information (e.g., the initial trust relationship and preference information) as in Part B Section VI. It is straightforward to observe that if the consensus threshold $\beta \le 0.9189$ (the initial global consensus degree), it means that the consensus threshold has been reached and there is no need to activate the feedback mechanism; otherwise, the subsequent CRP is required. Therefore, in order to explore the performance of the consensus model under different thresholds, we assign six values $\{0.92, 0.93, \ldots, 0.98\}$ to β successively. We show the performance of the proposed model in terms of three key indicators, namely, the number of iterations, the adjustment cost, and the final global consensus degree, in Table XVIII.

TABLE XVIII KEY INDICATORS OF THE CONSENSUS MODEL UNDER DIFFERENT THRESHOLDS

Consensus threshold	0.92	0.93	0.94	0.95	0.96	0.97	0.98
The number of iterations	1	2	3	4	5	6	9
GCD	0.9230	0.9312	0.9425	0.9575	0.9666	0.9731	0.9847
The adjustment cost	0.26	0.57	1.30	1.69	2.29	2.95	3.35

Clearly, with the increase of consensus threshold, the number of iterations gradually increases. This means that the higher the requirement for global consensus, the more times DMs need to conduct the modification. However, it is not difficult to find that no matter how the consensus threshold is increased, this model can effectively improve the global consensus degree in a limited number of iterations. Even when $\beta = 0.98$, it takes only 9 iterations to reach the consensus requirement. The observations demonstrate that the proposed model can efficiently help DMs reach a consensus and thus solve the LSGDM problem.

2) Sensitivity Analysis on Number of Clusters

The number of clusters is another important parameter that could influence the performance of the consensus model. As a rule of thumb, the number of clusters is usually set between 3 and 6. Therefore, in this section, we take four different cluster numbers and observe the performance of the model in terms of three key indicators, namely, the validity index *VI* of clustering, the number of iterations, and the adjustment cost. Note that since multiple iterations are involved in for each cluster number, we average values of VI derived from the iterations to facilitate the horizontal comparison among different cluster numbers. The experimental results are shown in Table XIX.

TABLE XIX KEY INDICATORS OF THE CONSENSUS MODEL UNDER DIFFERENT NUMBER OF CLUSTERS

The number of clusters	3	4	5	6
The average *VI*	1.05	1.10	1.14	1.15
The number of iterations	2	4	7	9
The adjustment cost	1.13	1.69	1.84	2.24

With the increasing number of clusters, the average *VI* of clustering also increases gradually. This means that when the number of clusters increases, more detailed and sophisticated structure of the DMs could be found. However, on the other hand, clusters can be regarded as participants in group-level decision-making. With more clusters, it becomes more difficult to reach the group-level consensus. In the case of the same consensus threshold, the larger the number of clusters, the more iterations and higher the adjustment cost it takes to reach consensus.

From the sensitivity analysis, the insight one may obtain is that in different practical problems, different indicators may be emphasized. For example, in emergency decision-making, rapid decision-making process is emphasized and more attention is paid to the time spent on iterations; in problems with complex social relationship among DMs, more attention is paid to the detailed clustering of DMs. Therefore, in view of different practical problems, one can refer to the above rules and determine the consensus threshold and the number of clusters according to actual needs.

D. Comparative Study

To further highlight the advantages of the proposed model, we sort out the characteristics of recent research and conduct a comparative analysis with them. The comparison is carried out in terms of the clustering process, the social relationship construction, and the feedback mechanism, as shown in Table XX.

We further analyze the difference between the proposed model with other studies in detail. In the clustering process, there are two core elements, namely the basis of clustering and the clustering algorithm. In most studies, preference similarity is used as the basis of clustering, that is, DMs with similar opinions are grouped into the same cluster. However, in LSGDM problems with a large number of DMs, such clustering is often not accurate enough. Since trust relationship can be the motivation and foundation of cooperation and has many advantages in solving decision-making problems, some studies also take trust relationship as the basis of clustering. However, most of the introduced clustering algorithms, such as *K*-means and Fuzzy *C*-Means, do not fully match the characteristics of trust-based clustering. In our model, we not only take the trust relationship as the basis of clustering, but also introduce spectral clustering algorithm considering the characteristics of the trust relationship. By cutting the trust network, DMs who trust each other are divided into the same cluster, which not only simplifies the LSGDM problem, but also enables DMs to communicate more smoothly, thus improving the efficiency of consensus reaching.

TABLE XX THE COMPARISON WITH RECENT LITERATURE

References	**The basis of clustering**	**Clustering algorithm**	**Social relationship**	**Object to conduct modification**
Tang et al. [21]	Preference similarity	Fuzzy *C*-means	×	DM/cluster
Liu et al. [7]	Preference similarity	Intuitionistic fuzzy *C*-means	×	DM
Wu & Xu [14]	Preference similarity	*K*-means	×	DM
Zhang et al. [18]	Preference similarity	Original algorithm	×	DM
Zha et al. [23]	Preference similarity	Fuzzy *C*-Means	×	Cluster
Ding et al. [25]	Trust relationship	Intuitionistic fuzzy *C*-means	Preference similarity	×
Zheng et al. [15]	Preference similarity compatibility	Original algorithm	×	DM
Lu et al. [2]	Trust relationship	*K*-means	Evaluation information	Cluster
Wu et al. [22]	×	Random	×	Cluster
Liu et al. [26]	×	×	Evaluation information	DM
Xu et al. [19]	Information consistent degree	Original algorithm	Evaluation information	Cluster
Zhang et al. [28]	Trust relationship	Two-stage trust network partition model	Evaluation information	DM
Our model	Trust relationship	Spectral clustering	Evaluation information & preference similarity	DM

It is observed that there are two common ways to construct trust relationship, that is formed either according to the subjective evaluation information given by DMs or according to the similarity of opinions between DMs. However, the trust relationship constructed by the former method is completely static, which is contrast to the fact that the trust relationship between DMs is constantly changing with the cooperation process. The latter approach, on the other hand, does not take into account that DMs may have already known each other and formed a certain relationship before participating in the decision-making process. Considering that neither of the methods is completely realistic, in our model, we take into account both the DMs' subjective evaluation information about each other and the preference similarity, and build a hybrid trust network which will be dynamically updated with the CRP.

To help DMs reach a consensus, most studies have introduced feedback mechanisms which typically involve (a) identifying the objects that need to conduct modification; and (b) generating modification suggestions for these objects. For LSGDM problems, it is often costly and time-consuming to directly identify individual DMs for modification. Therefore, some studies with clustering process take clusters as the operation objects, and directly modify the collective preference of clusters. However, in this way, it directly ignores whether the DMs in the cluster are willing to make changes, which may cause some DMs to feel that their opinions are distorted and ultimately oppose to the decision results. Therefore, in our model, we first identify the clusters that need to conduct modification and generate modification suggestions for them with the goal of maximizing the global consensus. We do not require to modify the cluster collective preference directly, but feedback the suggestions to DMs in the cluster for a reference. On this basis, by minimizing the adjustment cost, the modification suggestions are generated for the individual DMs in the cluster to balance the consensus efficiency and the willingness to cooperate of the DMs.