Information Sciences 575 (2021) 499–527

Contents lists available at ScienceDirect

Information Sciences

journal homepage: www.elsevier.com/locate/ins

An efficient consensus reaching framework for large-scale social network group decision making and its application in urban resettlement

Xiangrui Chao [a], Gang Kou [b], Yi Peng [c,*], Enrique Herrera-Viedma [d,e], Francisco Herrera [d,f]

[a] Business School, Sichuan University, Chengdu 610064, China
[b] School of Business Administration, Southwestern University of Finance and Economics, Chengdu 611130, China
[c] School of Economics and Management, University of Electronic Science and Technology of China, Chengdu 611731, China
[d] Andalusian Research Institute in Data Science and Computational Intelligence, Dept. of Computer Science and AI, University of Granada, Granada 18071, Spain
[e] Department of Electrical and Computer Engineering, Faculty of Engineering, King Abdulaziz University, Jeddah 21589, Saudi Arabia
[f] Faculty of Computing and Information Technology, King Abdulaziz University, Jeddah 21589, Saudi Arabia

A R T I C L E I N F O

Article history:
Received 24 March 2020
Received in revised form 12 June 2021
Accepted 15 June 2021
Available online 18 June 2021

Keywords:
Large-scale decision-making
Consensus reaching
Dual social network
Preference classification
Minimum consensus cost

A B S T R A C T

Urban resettlement projects involve a large number of stakeholders and impose tremendous cost. Developing resettlement plans and reaching an agreement amongst stakeholders about resettlement plans at a reasonable cost are some of the key issues in urban resettlement. From this perspective, urban resettlement is a typical large-scale group decision-making (GDM) problem, which is challenging because of the scale of participants and the requirement of high consensus levels. Observing that residents who are affected by a resettlement project often have tight social connections, this study proposes a framework to improve the consensus reaching and uses the minimum consensus cost to reduce the total cost for urban resettlement projects with more than 1000 participants. Firstly, we construct a network topology that consists of two layers to deal with incomplete social relationships amongst large-scale participants. An inner layer consists of participants whose preference similarities and trust relations are known. Meanwhile, an outside layer includes participants whose trust relations cannot be determined. Secondly, we develop a classification method to classify participants into small subgroups based on their preference similarities. We can then connect the participants whose trust relations are unknown (the outside layer) with the ones in the inner layer using the classification results. To facilitate effective consensus reaching in large-scale social network GDM, we develop a three-step approach to reconcile conflicting preferences and accelerate the consensus process at the minimum cost. A real-life urban resettlement example is used to validate the proposed approach. Results show that the proposed approach can reduce the total consensus cost compared with the other two practices used in the actual urban resettlement operations.

1. Introduction

Urban resettlement is related to regional economic development, improvement of residents' living conditions, social stability and ending poverty [1,25]. The successful implementation of a resettlement project depends on many factors, such as

* Corresponding author.
E-mail address: pengyi@uestc.edu.cn (Y. Peng).

https://doi.org/10.1016/j.ins.2021.06.047
0020-0255/© 2021 The Author(s). Published by Elsevier Inc.

politics, economics, sociology and law [39]. Developing resettlement plans and reaching an agreement amongst households about resettlement plans at a reasonable cost are some of the key issues in urban resettlement. Households have various knowledge backgrounds and social status and often have tight social connections. From this perspective, urban resettlement is a large-scale social network group decision-making (LSSNGDM) problem, which is challenging because of the scale of the participants and the requirement of high consensus levels.

In LSSNGDM, the social network relation includes trust relation and preference similarity relation [12]. The trust relation is a Boolean function, which assumes the non-existence or existence of a trust relationship between two decision makers (DMs). The preference similarity relation measures the similarity of two DMs' preferences for a set of alternatives. Based on the above two network relationships, a consensus reaching process consists of three steps: (1) build a social network of DMs. A commonly used method constructs a social network based on trust relationships between DMs [12,35]. Wu et al. [38] built a social network by using a trust degree calculation and a trust-based preference collection model. Another way to build a social network is to utilise the preference similarities between DMs [24,41]. (2) Divide a large group of DMs into smaller subgroups by using social network analysis. Dividing a large group of DMs into several subgroups is an effective tool to reach a consensus in LSSNGDM [41]. Given that DMs in a subgroup have similar preferences and trust each other, they are more likely to change their preferences to reach a consensus than DMs in a larger group. The mainstream methods include clustering DMs based on individual preference relations [35] and classification based on opinion dynamics [12]. (3) Construct consensus models for each subgroup. The consensus reaching mechanism in social network GDM is similar to traditional GDM [12,35,41]. The most widely used method is the interactive consensus reaching framework, in which DMs are asked to modify their preferences by using feedback information until a consensus is reached [7,19].

The consensus reaching in LSSNGDM is more complex than that in small size social network GDM due to the following reasons. Firstly, heterogeneous preferences exist in GDM because of DMs' diverse educational backgrounds, knowledge and decision habits [14,34,38]. Methods for measuring the distance amongst heterogeneous preferences must be developed for classification purpose. Urena et al. [24] analysed the network structure on the basis of the similarity degree between heterogeneous preferences and proposed a consensus model on the basis of the confidence degree. Secondly, the trust relations are partially missing, and the network structures of LSSNGDM are unknown [12,33,35]. Wu et al. [37] employed a trust propagation method to visualise the consensus reaching process. Wu et al. [40] proposed a shortest path algorithm to divide the network to reach a consensus for LSSNGDM. Lastly, traditional methods, such as the feedback adjustment mechanism and DMs' weight determination methods, are inefficient when dealing with LSSNGDM problems [35,41] which require new methods to improve the efficiency of the consensus reaching process [23]. Cheng et al. [7] developed a minimum cost consensus model for 20 DMs by using the social network information to determine the weights.

Although many consensus models have been developed to handle less than 200 experts [28.36.42.43] studies on large-scale (with more than 1000 participants) social network structure, opinion classification and consensus reaching have been scant [9]. Existing research is difficult to tailor to meet the needs of LSSNGDM problems. Firstly, the existing studies construct social networks on the basis of the trust relations or similar preferences amongst DMs [12,34,35,37,38,41]. In real-life LSSNGDM problems, a DM is influenced not only by people he/she trusts but also by people with similar preferences. For example, in the urban resettlement project studied in this study, the residents with similar preferences and trust relations are more likely to reach an agreement. Secondly, the commonly used feedback adjustment mechanisms involve multiple rounds of negotiations, which largely increase the time and human resource costs when the number of DMs is large [23,34,41]. This study takes the urban resettlement project as an example. The real estate developer hired 62 employees, including seven senior managers, and took approximately 6 months to complete the household survey and discussions. The salary expenditure alone cost approximately 1.02 million-yuan RMB. Gong et al. [16] proposed a minimum cost consensus reaching model to improve the feedback adjustment consensus mechanism for small size GDM problems (five experts). This study is different from [16] because we consider large-scale (with more than 1000 participants) social network relationships between DMs. In addition, their model was designed for crisp value and cannot be directly used for the preference relation (matrix or vector). Lastly, trust relations between DMs are important information in the consensus-reaching process in LSSNGDM. In real-life practice, only a small part of the trust relations amongst a limited number of DMs can be established due to the resource constrains. If the trust relationship is incomplete, then a special consensus-building model needs to be developed [12,17]. The existing studies use either the transitivity of the trust paths to predict the trust relations [12,40,41] or the preference relationships to replace the trust relations [34,35]. Gupta [17] proposed to achieve consensus without filling in the trust relations. They used the influence intensity of DMs, whose trust relationships were unknown, to determine the possibility of change in their opinions and established a group consensus. The existing methods are not effective in filling trust relations and reaching consensus for large-scale social networks. Thus, new approaches, such as classification algorithms, are needed to improve the efficiency of consensus reaching in LSSNGDM.

This study aims to develop a consensus reaching framework for LSSNGDM (with more than 1000 participants) by using analytical tools (social network analysis, data mining and optimisation) to achieve an automatic consensus building with efficiency and controlled cost. To achieve this goal, we firstly propose a general topological structure to address the partially missing trust relation problem by connecting DMs with complete trust relation DMs who do not have clear trust relations. Then, we develop a classification method to divide a large-scale social network into small subgroups on the basis of individual preference relations. Thus, all DMs' preferences can be reflected in the GDM process. Lastly, we build an optimisation model to facilitate preference modifications, which replaces the repeated iterations and obtains an optimal consensus cost. We apply the proposed approach to a real-life large-scale urban resettlement project to validate it. The results show that the proposed approach can improve the efficiency of consensus reaching of large-scale urban resettlement and reduce the total consensus reaching cost.

This study contributes to the literature in two ways. Firstly, we develop a two-layer suspension social network structure to fuse trust relation and preference similarity for large-scale decision-making participants. We propose a classification method by using support vector machines (SVM) in vector space to deal with missing trust relations. Secondly, prior work on LSSNGDM has mainly focused on problems with less than 200 participants. In addition, consensus cost is a crucial issue in many real-life LSSNGDM projects. However, no current research in LSSNGDM has taken the consensus cost into consideration. To the best of our knowledge, our study is one of the first few to optimise the consensus cost of a real-life LSSNGDM project with 1861 participants. The minimum cost optimisation model constructed in this study enriches the toolbox of consensus reaching in LSSNGDM problems.

The remainder of this paper is organised as follows. Section 2 introduces the basic features of LSSNGDM and outlines the proposed three-step consensus reaching framework. Section 3 describes the first step in the proposed framework: how to construct a social network structure. Section 4 describes the second-step: a classification approach to divide DMs into smaller groups. Section 5 introduces the last step: develop a consensus-reaching process with minimum compensation cost. Section 6 uses a real-life urban resettlement project to validate the usefulness and effectiveness of the proposed approach. Section 7 concludes the paper.

2. Literature review and preliminaries

A consensus reaching process [19] is implemented to improve the consensus degree of LSSNGDM by dividing large-scale groups into smaller subgroups and reaching satisfactory consensus in subgroups [35]. This section introduces prior literature on the basic features of LSSNGDM and the components of a general consensus reaching framework in LSSNGDM and describes the consensus reaching framework proposed in this study.

2.1. Basis features of LSSNGDM

The LSSNGDM has three prominent features (Fig. 1): a large number of DMs, heterogeneous preference formats used by DMs and interactions between DMs [12].

A GDM problem is normally regarded as a large-scale GDM problem when the number of participants is larger than 20 [24,40,43]. The largest number of DMs that has been studied in the existing literature is less than 200.

Given that DMs use diverse preference expressions in LSSNGDM [12] the consensus reaching process under the heterogeneous preferences environment has been an important research topic in large-scale GDM [14,19] and social network GDM (SNGDM) [34]. The common preference structures include preference ordering, preference value, multiplicative preference relation and additive preference relation [22,33,34]. Appendix A introduces the four most frequently used preference representation forms.

In a GDM problem, DMs in the same social network share experience and knowledge and may influence each other's viewpoints by either referring to others' opinions or trusting neighbours [37,38]. Urban resettlement is a typical LSSNGDM problem that involves thousands of residents with tight social connections in the consensus reaching process. A consensus is difficult to reach in urban resettlement projects due to conflicting preferences and the requirement of reaching high consensus levels at a reasonable cost.

2.2. General consensus reaching process for LSSNGDM

We propose a three-step consensus reaching framework to facilitate consensus reaching in LSSNGDM problems (Fig. 2). The first step constructs a social network structure for a LSSNGDM problem, which will be described in Section 3. The second step divides the DMs into smaller subgroups. The last step utilises a consensus reaching process to efficiently reach consensus with minimum cost. The following subsections outline the second and third steps.

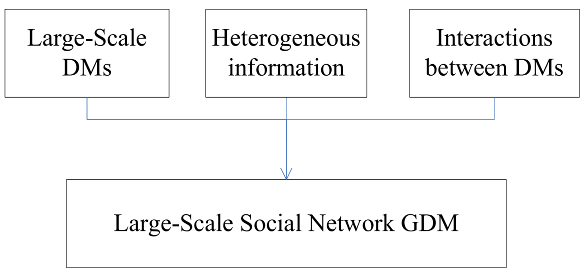

Fig. 1. Three basic features of LSSNGDM.

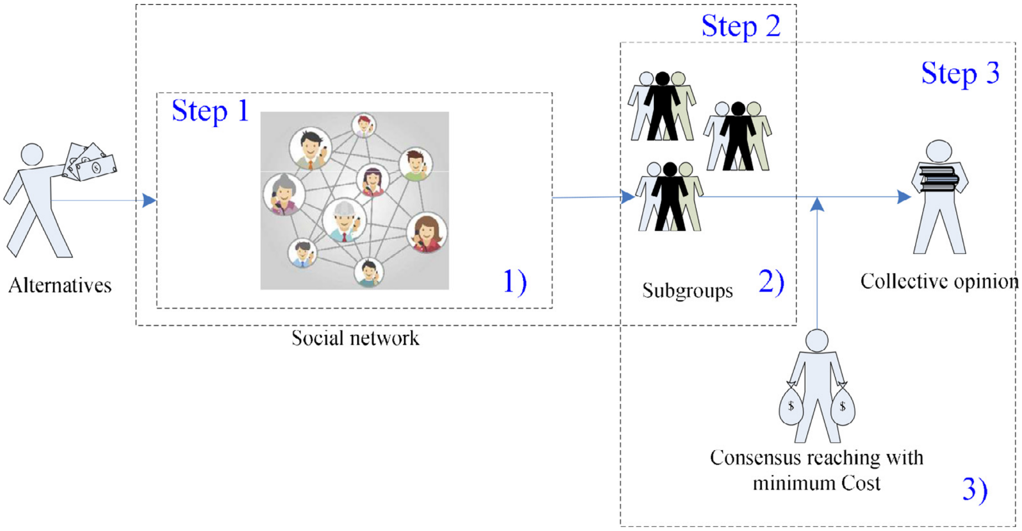

Fig. 2. Proposed LSSNGDM consensus reaching framework.

2.2.1. Subgroup partition

Subgroup partition is used to improve the efficiency of large GDM. Clustering and classification are the commonly used methods to partition subgroups. Palomares et al. [27] proposed a fuzzy clustering and outlier detection method to cluster DMs to small groups according to their preferences. Wu and Xu [42] proposed a changeable cluster algorithm to capture preference evolution. Wu et al. [40] developed a two-stage subgroup partition algorithm to reduce the complexity of large-scale GDM problems. Urena et al. [33] proposed to use the degree of self-confidence as a criterion to group DMs. In the above-mentioned methods, the subgroups are partitioned through either clustering or the features of network structures. They require complete network structures and DMs' information, such as trust relation, self-confidence degree and support degree of opinions. Furthermore, they need repeated adjustments of individual preferences to reach a high consensus degree, which is inefficient for large GDM.

Classification algorithms divide a large number of participants into small subgroups. In LSSNGDM, the labels for subgroup classification are determined on the basis of the social relationships amongst DMs. Given that DMs in a subgroup have similar preferences or trust each other, we can divide large-scale DMs into subgroups by using preference or trust relations. We can identify preference and trust relations in a social network through interviews and surveys with DMs and assign labels to subgroups, which are used to train classifiers.

The application of classification algorithms in subgroup partition requires two necessary conditions. Firstly, the training data must have classification labels, which can be obtained on the basis of the management experiences, background knowledge, interviews and questionnaires. The second condition is that the data should be transformed into a matrix, in which the columns are attributes, and the rows are data records. Traditional classifiers cannot be directly used in LSSNGDM because a preference relation is a matrix, and the entire matrix must be converted into a vector for training and prediction. Moreover, classification labels are difficult to determine in large-scale social networks. This study proposes a classification approach, which is described in Section 4, to handle these issues in LSSNGDM.

2.2.2. Consensus reaching in LSSNGDM

The traditional feedback mechanism of asking each DM to repeatedly modify their preferences to reach an agreement is inapplicable when the number of DMs is large. The two major directions of applying social network analysis to reach consensus in LSSNGDM are as follows. The first one is to build a trust relation to reach consensus within subgroups; the other one is to use opinion dynamics to improve consensus reaching, which is based on the evolution of individual DM's opinions or preferences in a social network [12,34,35,42].

Time and cost are two important issues in LSSNGDM. Considerable time and resources are needed to investigate the preference and trust relations amongst DMs. In many real-life GDM problems, management or related institutions compensate participants to change their opinions, which is called consensus cost. Furthermore, the complex structures of social networks demand more human resources and costs to detect conflicts amongst DMs and coordinate their opinions. Labella et al. [23] claimed that the time cost optimisation in the consensus process is a key issue because the traditional GDM models take more time to converge when the number of DMs increases. Zhang et al. [46] analysed the efficiency of different consensus processes and found that the interactive feedback adjustment mechanisms are often inefficient. Zhang et al. [45] proposed a minimum cost consensus model under a soft consensus condition; however, their research did not consider social network amongst DMs. Cheng et al. [7] developed a minimum cost consensus model in social network GDM, and only 20 DMs were involved. Many studies have assigned DMs with different weights to reflect their importance, such as the piecewise function [43] matrix method [13] and local and whole weights modification [27] to cost-effectively and efficiently reach a consensus.

Previous work has shown that the amount of compensations required to change DMs' preferences varies [6,16] which implies that economic compensations can be used to avoid or reduce the number of repeated iterations of preference modifications. One possibility is to seek an equilibrium that balances the costs and benefits of DMs [6,16,20]. Zhang et al. [45] showed that minimum consensus cost models can reduce the total cost and improve decision efficiency. In Section 5.2, we develop an optimisation model to determine the optimal consensus cost for LSSNGDM.

Section 3 describes the first step of the proposed consensus reaching framework in detail.

Table 1 summarises the notations used in the rest of this paper.

3. Social network structure of LSSNGDM

In LSSNGDM, participants are connected through social relations. Establishing a social network structure helps in understanding the associations between participants and handling incomplete information. This section proposes an approach to construct a network structure for LSSNGDM, detect different communities and identify their importance in a social network.

3.1. Social network analysis in GDM

A social network consists of a set of nodes, which indicate individuals or organisations, and dyadic ties, which represent social relations and interactions amongst the nodes of the network [32]. Social network has been widely used to reveal the structures and patterns of a social group [9,32]. Social relations that connect DMs are the basic components of a social network. Individuals located at the centre of a social network have more influence in the consensus building process than others because they have more connections in the social network [10,12,35]. Existing studies [10,12,35] construct social networks of GDM by using mutual trust relationships or preference similarity relations between DMs. In practical LSSNGDM applications, the complete trust relationships between DMs are difficult to fully construct due to missing information and complicated social relations.

This study aims to address this gap by developing a local network structure to establish social networks for LSSNGDM with incomplete trust and similarity relations. The first step identifies DMs whose trust relationships and similarity relationships are known and builds a *local network* with complete trust relations and similarity relationships amongst DMs. In the second step, DMs whose trust relationships are unknown are divided into different classes on the basis of the preference similarity relations between the external points and the points in the local network.

A social network can be illustrated using a graphic, algebraic or sociometric form. For example, the relation of n DMs in a social network between any two DMs can be presented as a sociometric form that is an adjacent matrix with the following function:

$$A = (e_{ij}) = \begin{cases} e_{ij} = 1, & if \quad e_i R e_j \\ e_{ij} = 0, & if \quad e_i \ R^{-} \ e_j \end{cases} \tag{1}$$

where $e_i R e_j$ means that two DMs have a connection, which is either a preference similarity relation or a trust relation. Otherwise, the entry is zero. $A = (e_{ij})$ is also an adjacent matrix.

Table 1
Notations used in the consensus reaching framework.

Symbols	Descriptions	Symbols	Descriptions
X	Alternative set	N	Number of *DM*
M_v	Community of social network (Section 3.3.2)	m_v	Number of suspend vertexes that belong to in M_v
DM_v	DMs in each community	s_v	Weighted centrality degree(Definition 1)
$\rightarrow$	Vector	ω	Derived priority vector of preference relation
$S(\vec{p}_i, \vec{p}_j)$	Similarity measure of two preference vectors	U	Utility value
<,>	Inner products of two vectors	O	Preference ordering
A	Adjacent matrix	P	Multiplicative preference relation
e_{ij}	Entry of A	B	Additive preference relation
A'	Weighted adjacent matrix (vector)	u_i	Entry of the utility value
f_i	Entry ofA'	o_i	Entry of preference ordering
D'	Distance matrix	a_{ij}	Entry of multiplicative preference relation
d'_{ij}	Entry ofD'	b_{ij}	Entry of additive preference relation
D	Weighted distance matrix	u'_k	$u_k / \sum_i u_i$
d_i	Entry of D	o'_k	$\frac{n-o_k}{n-1} / \sum_{i=1}^{n} \frac{n-o_i}{n-1}$
M	Dual social network (Fig. 4)	$DS(,)$	Distance measurement of different preference relations
$\hat{m}_{ij}$	Entry of dual network	o_C	Collective group preference
v	Vertex of M	$K\left(\vec{r}_i, \vec{r}_j\right)$	Kernel function in SVM

3.2. Local social networks in LSSNGDM

Mutual influence and group behaviour dominate the consensus building process in LSSNGDM [28,29,38]. The different preference classes and the trust and preference similarity relations amongst DMs must be understood to efficiently reach a high degree of consensus in social networks.

Although identifying all the trust relations of DMs in a large-scale social network is impossible due to privacy protection and lack of information, a local social network (such as representatives of DMs) with its relative complete nodes and links can be determined through questionnaires, statistics and interviews.

3.2.1. Proposed topological structure for LSSNGDM

Fig. 3 shows the proposed general social network structure. In this structure, the nodes are DMs, and the links are the relationships amongst DMs, which include trust relation and preference similarity. The inner layer (inside the red circle) is a complete local social network, in which the trust relationships between DMs are known, and the outside suspension nodes are the DMs whose trust relations with the inner nodes cannot be determined. We need to investigate the preferences of the outside nodes and the nodes in the local social network to connect the former to the local social network (Section 4).

3.2.2. Social relations in a local social network

This section investigates the inner structure of a local social network to formulate different preference classes. Existing research on social network of GDM only established a single network, in which a link between two DMs is either a trust relation or a preference similarity [12,34,37,38]. In many real-life decision-making situations, a DM is influenced not only by people he/she trusts but also by people with similar preferences. Two DMs with similar preferences and no trust relation are more likely to reach an agreement than those have trust relation and opposite preferences. Thus, trust relation and preference similarity should be used in classifying DMs.

On the basis of this observation, we propose a dual structure of a local social network (Fig. 4), which considers preference similarity and trust relation. In Fig. 4, the base layer is a connection relation in a social network graph (for instance, trust relation), and the embedded layer is a preference similarity relation between nodes. The base layer is embedded in the network established on the basis of the preference similarity degree.

In real-life decision problems, the base layer is constructed using the trust relations obtained from investigations, questionnaires and data analysis. The adjacent relation R is determined by mixed indexes, such as a distance in the social net-

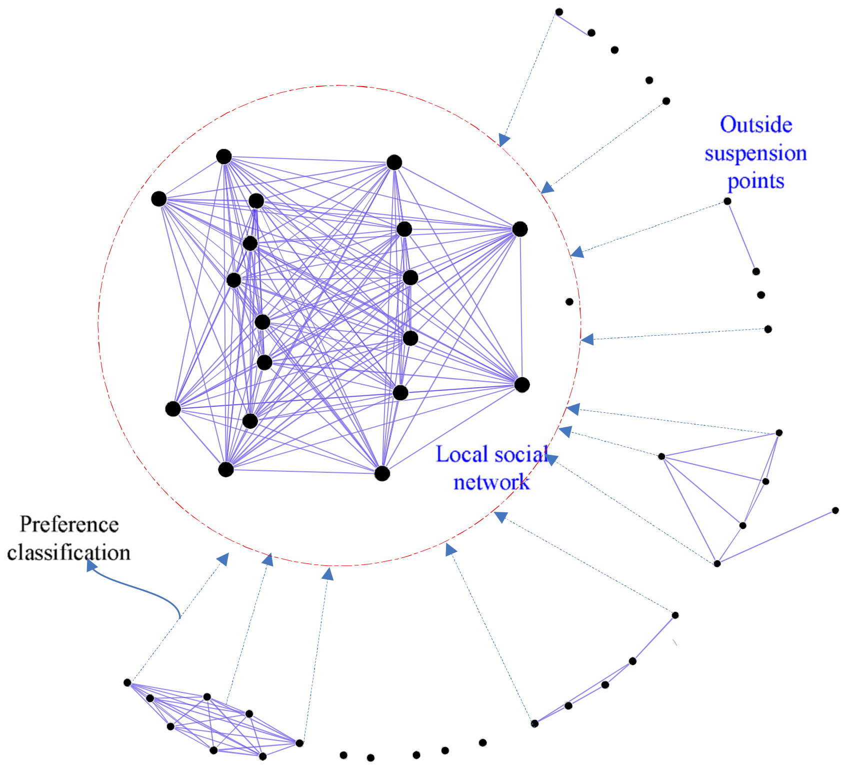

Fig. 3. Topological structure of LSSNGDM.

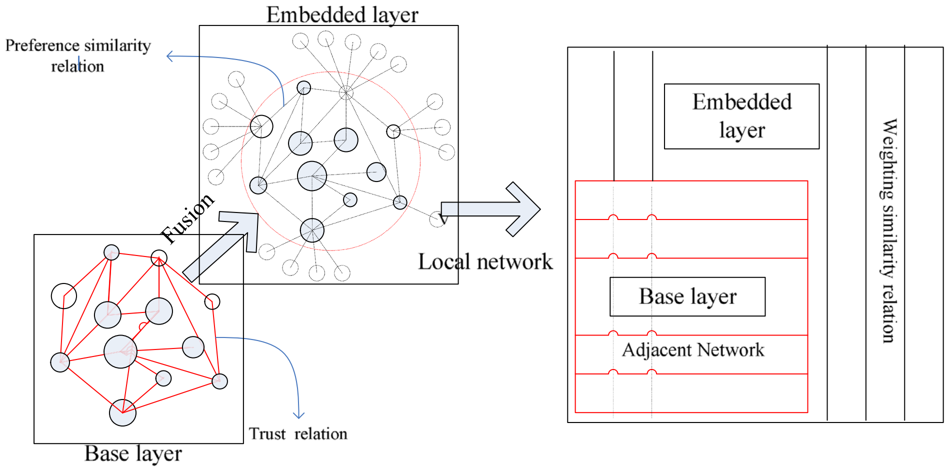

Fig. 4. Dual structured network of local social network.

work, common friends and intimacy of the relationship [29]. The embedded layer is built on the basis of the similarities amongst DMs, which are measured using individual preference relations.

The outside suspension nodes in Fig. 4 are connected with the inside nodes through the preference similarity relations amongst the outside nodes and the nodes in the local network. The size of outside nodes is proportional to its influence in a LSSNGDM problem.

3.2.3. Fusion of the trust relation and preference similarity relation in a local social network

The construction of the dual structured network follows three steps:

Step 1 Construction of the dual structure of a social network

1) The base layer (Fig. 4) is a network composed of nodes and links. In this network, the nodes are DMs, and the links are trust relations. We use A to note an adjacent matrix representing trust relations.

2) The embedded layer (Fig. 4) is also a network, in which each link represents a preference similarity. In urban resettlement, people with similar social status, such as household earnings and family structures, are likely to have similar preferences and make similar decisions. This study assumes that p_i and p_j are two preference relations in a local social network provided by DM_i and DM_j, respectively. The distance measure based on similarity degree is defined as follows:

$$d'_{ij} = S(p_i, p_j) \tag{2}$$

where $S(p_i, p_j)$ is the cosine similarity measure of two preference vectors.

Step 2: Weighted vectorisation of the dual network

Let pv_i be the number of suspension nodes attached to node v_i in a local social network. Weighted vector C (belonging to$\Re^{n\times1}$) is a normalised vector and an entry $c_i = pv_i/\sum_k pv_k$, where k is the number of nodes in the local social network. Each entry in this vector C is the ratio of DMs who have a similar preference with node v_i. Step 2 uses weighted vector C to absorb the influence of outside nodes on consensus reaching into the local social network. Then, weighted adjacent matrix A' can be formulated by $A' = (f_i) = AC^T$ at the base layer, wheref_i is an entry of weighted adjacent matrix (vector) A', and A is an adjacent matrix. The distance matrix is $D' = (d'_{ij})$, and the weighted distance matrix is $D = (d_i) = D'C^T$.

Step 3: Unitisation and fusion

This step unitises the elements of each weighted adjacent matrix and weighted distance matrix by their standard deviation:

$$\widehat{f}_i = \frac{f_i - \frac{1}{n}\sum_i f_i}{\sqrt{\frac{1}{n-1}\sum\limits_i \left(f_i - \frac{1}{n}\sum\limits_i f_i\right)^2}} \tag{3}$$

$$\hat{d}_i = \frac{d_i - \frac{1}{n}\sum_i d_i}{\sqrt{\frac{1}{n-1}\sum\limits_i \left(d_i - \frac{1}{n}\sum\limits_i d_i\right)^2}} \tag{4}$$

The following matrix (5) combines the weighted adjacent matrix and weighted distance matrix by using weight θ to fuse the above-mentioned two layers. Let H be a new matrix with two sub-blocks:

$$H = \left[A'_{n\times 1} | D_{n\times 1}\right] \tag{5}$$

where '|' represents a block matrix. Then, we can use the singular value decomposition of H with matrixes $U_{n\times 2}$ and $V_{2\times n}$ to determine the weights of two separate networks with the following formula:

$$H = U\begin{bmatrix} \bar{\theta}_1 & \cdot \\ \cdot & \bar{\theta}_2 \end{bmatrix} V \tag{6}$$

Finally, the weights are computed as $\tilde{\theta} = \bar{\theta}_1 / \sum \bar{\theta}_i$ and $1 - \tilde{\theta} = \bar{\theta}_2 / \sum \bar{\theta}_i$. $\hat{m}_{ij}$ represents the combination of the base and embedded layers:

$$\hat{m}_{ij} = \tilde{\theta} e_{ij} + (1 - \tilde{\theta}) d'_{ij} \tag{7}$$

where e_{ij} and d'_{ij} are entry of the adjacent matrixes and distance matrix between DMs i and j, respectively. Parameter θ combines e_{ij} and d'_{ij} to form the composed link $\hat{m}_{ij}$.

3.2.4. Centrality of the local social network

We define a weighted eigenvector centrality to determine the centrality of each node in a dual social network (Definition 1). The idea is to detect the nodes that have the most connections to the other nodes in a network. In many cases, the centrality reflects the influence of a DM in a social network. A node with a higher centrality has more influence because it has more connected links with other nodes.

Definition 1. Weighted eigenvector centrality: In a given dual social network, $M = (\hat{m}_{ij})$ is a combination of the base and embedded layers. The weighted eigenvector centrality of node v is defined as follows:

$$s_v = \frac{1}{\sigma}\sum_{j=1}^{n} \hat{m}_{tj} v_t, \tag{8}$$

where σ is a constant that is computed as the largest eigenvector of M.

The proposed centrality score s_v can be used to minimise the induced modification cost. Decision makers with larger centrality scores have greater influence on the collective opinion, and they have higher weights in a decision-making process. Thus, their preference modification range is set smaller, and the corresponding consensus cost is lower.

4. Group classification in LSSNGDM

Classification is introduced to divide DMs into smaller groups to improve the efficiency of group negotiation and consensus reaching in LSSNGDM. To do so, we assign labels to different communities in the local social network on the basis of their complete trust relations. Then, we train a classifier and assign the outside nodes, whose complete trust relations are unknown, to each class to ensure that all DMs can be divided into different subgroups. Although the data used in traditional classifications are vectors, some preference relations (such as multiplicative and additive) are matrixes. Therefore, the heterogeneous preference relations, which often occur in LSSNGDM, should be transformed into a standard classification vector with labelled classes and vectorisation. To deal with heterogeneous preference relations that often occur in LSSNGDM, we transform them into a uniform data format by using similarity measure. Appendix A provides the theoretical basis for this transformation, which is an inner product space composed of different preference relations. This study chooses SVM as the classifier because it is one of the widely used classification methods and is suitable for preference relation data, which has a small number of dimensions and a large volume of data [31].

4.1. Inner product space

We use inner product space, instead of Euclidean space, to determine the 'distances' amongst different preference relations due to the diverse structures of different preference relations.

The commonly used preference relations include utility value, preference ordering, multiplicative preference relation and additive preference relation. Kou and Lin [21] and Chao et al. [2] developed the cosine similarity relation between these preference relations and their derived vectors. However, the cosine similarity degree measure is not a mathematical distance measure (does not satisfy the triangle inequality). We transform each preference into a unitised vector to make it suitable for classification, and the cosine similarity degree of two preference relations becomes an inner product. The introduction of the different preference relations and the detailed transformation process are summarised in Appendix A.

Definition 2. The unitized vector of utility value is $\vec{u}$, the preference ordering is $\vec{o}$, the unitized column vector of the multiplicative preference relation is $\vec{a}_j$, and the additive preference relation is $\vec{p}_j$. The $< \cdot, \cdot >$ is an inner product between two vectors. Then, the distance measurement DS amongst different preference relations is computed as follows:

Case 1: The inner product between column vectors (utility values and preference ordering) is defined as follows:

$$DS(U,O) = \frac{1}{n}\sum_{j=1}^{n} < \vec{u}_j, \vec{o}_j >= \frac{1}{n}\sum_{j=1}^{n}\sum_{h=1}^{n}\bar{u}_{hj}\bar{o}_{hj} = \frac{1}{n}\sum_{j=1}^{n}\sum_{h=1}^{n}\frac{u_{hj}}{\sqrt{\sum_{i=1}^{n}(u_{ij})^2}}\frac{o_{hj}}{\sqrt{\sum_{i=1}^{n}(o_{ij})^2}} \tag{9}$$

Case 2: The inner product between matrixes (multiplicative and additive preference relation) is defined as follows:

$$DS(A,B) = \frac{1}{n}\sum_{j=1}^{n} < \vec{a}_j, \vec{p}_j >= \frac{1}{n}\sum_{j=1}^{n}\sum_{h=1}^{n}\bar{a}_{hj}\bar{p}_{hj} = \frac{1}{n}\sum_{j=1}^{n}\sum_{h=1}^{n}\frac{a_{hj}}{\sqrt{\sum_{i=1}^{n}(a_{ij})^2}}\frac{p_{hj}}{\sqrt{\sum_{i=1}^{n}(p_{ij})^2}} \tag{10}$$

Case 3: The inner product between a column vector and a matrix is defined as follows:

$$\begin{aligned} DS(U,MU) &= \tfrac{1}{n}\sum_{j=1}^{n} < \vec{u}_j, \vec{a}_j >= \tfrac{1}{n}\sum_{j=1}^{n}\sum_{h=1}^{n}\bar{u}_{hj}\bar{a}_{hj} = \tfrac{1}{n}\sum_{j=1}^{n}\sum_{h=1}^{n}\frac{u_{hj}}{\sqrt{\sum_{i=1}^{n}(u_{ij})^2}}\frac{a_{hj}}{\sqrt{\sum_{i=1}^{n}(a_{ij})^2}}, \\ DS(O,MU) &= \tfrac{1}{n}\sum_{j=1}^{n} < \vec{o}_j, \vec{a}_j >= \tfrac{1}{n}\sum_{j=1}^{n}\sum_{h=1}^{n}\bar{o}_{hj}\bar{a}_{hj} = \tfrac{1}{n}\sum_{j=1}^{n}\sum_{h=1}^{n}\frac{o_{hj}}{\sqrt{\sum_{i=1}^{n}(o_{ij})^2}}\frac{a_{hj}}{\sqrt{\sum_{i=1}^{n}(a_{ij})^2}}, \end{aligned} \tag{11}$$

$$\begin{aligned} DS(U,FU) &= \tfrac{1}{n}\sum_{j=1}^{n} < \vec{u}_j, \vec{p}_j >= \tfrac{1}{n}\sum_{j=1}^{n}\sum_{h=1}^{n}\bar{u}_{hj}\bar{p}_{hj} = \tfrac{1}{n}\sum_{j=1}^{n}\sum_{h=1}^{n}\frac{u_{hj}}{\sqrt{\sum_{i=1}^{n}(u_{ij})^2}}\frac{p_{hj}}{\sqrt{\sum_{i=1}^{n}(p_{ij})^2}}, \\ DS(O,FU) &= \tfrac{1}{n}\sum_{j=1}^{n} < \vec{o}_j, \vec{p}_j >= \tfrac{1}{n}\sum_{j=1}^{n}\sum_{h=1}^{n}\bar{o}_{hj}\bar{p}_{hj} = \tfrac{1}{n}\sum_{j=1}^{n}\sum_{h=1}^{n}\frac{o_{hj}}{\sqrt{\sum_{i=1}^{n}(o_{ij})^2}}\frac{p_{hj}}{\sqrt{\sum_{i=1}^{n}(p_{ij})^2}}. \end{aligned} \tag{12}$$

Moreover, it is easy to prove that $\frac{1}{n}\sum_{j=1}^{n} < \vec{\varpi}, \vec{p}_j >=< \vec{\varpi}, \frac{1}{n}\sum_{j=1}^{n}\vec{p}_j >$ and $\frac{1}{n}\sum_{j=1}^{n} < \vec{\varpi}, \vec{a}_j >=< \vec{\varpi}, \frac{1}{n}\sum_{j=1}^{n}\vec{a}_j >$ holds, where $\vec{\varpi}$ is a column vector. Thus, the preference relation with a matrix structure can be transformed into a column vector by row weighted algebraic mean and used to classify data.

According to the above-mentioned definitions, the transformed preference relations can be constructed as an inner product metric space (Appendix A).

4.2. Label assignment

In classification tasks, labels are predetermined from historical data and used to train classifiers. In this study, DMs in a community are assigned with the same class label, which is used to classify the outside suspension nodes by using SVM. DMs with the same label belong to one subgroup, which means that DMs in a subgroup may come from the same community or may be outside suspension nodes. The definitions of community and subgroup in this study are as follows:

Definition 3:. Community in a local social network. In a local social network, a set of nodes is called a community if the nodes are internally densely connected, which are measured by the denseness of the links between nodes.

Definition 4:. Subgroup in a social network. A community and outside suspension nodes, which are connected to the nodes in this community through classification, constitute a subgroup.

In this study, we divide a large-scale group into several subgroups. The following paragraphs introduce the procedure of community detection and subgroup classification.

Community detection methods [11,15] can be used to divide a social network into smaller parts, in which the nodes have a higher connection degree with each other (i.e. have more links in a network). Community detection methods are used to decrease the number of class labels to deal with the large number of nodes and improve efficiency in LSSNGDM. We employ a spectral analysis to detect different communities in a social network [11,15]. Fig. 5 shows the process of community detection.

Spectral analysis is a well-known community detection method [11,15]. It is a clustering method based on the Laplacian matrix, which is constructed using the network degree matrix and the correlation matrix. The Laplacian matrix constructed in this study is introduced as follows:

Firstly, we construct a Laplacian matrix (L) transformed from a similarity matrix (S) of an inner product of the preference relations amongst the DMs. Then, the K-means clustering algorithm based on a Laplacian matrix is implemented to identify communities in the local social network.

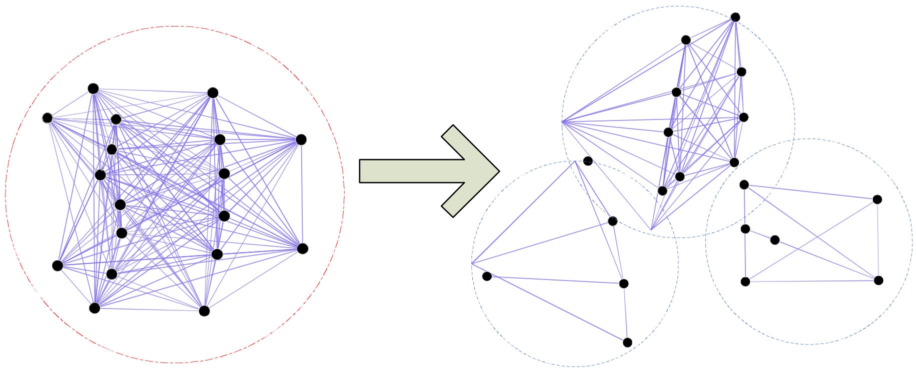

Fig. 5. Community detection [11].

$$L = D^{-1/2} S D^{-1/2} \tag{13}$$

where D is the degree matrix, and$S = (s_{ij})$, in which $s_{ij} = e^{-\frac{(<p_i, p_j>)^2}{2\sigma^2}}$, and p_i, p_j are preference relations of DM_i, DM_j.

4.3. Inner product SVM

Preference relations in LSSNGDM are characterised by a small number of features and a large volume of data. SVM, which discovers an optimal hyperplane to separate data into different classes, is suitable to classify this type of data. The traditional Euclidean space SVM is no longer applicable because DMs' preferences are often heterogeneous in large-scale GDM. This study develops an SVM on the inner product space for LSSNGDM. We need to build an optimisation model in the inner product space to determine the support vectors in an inner product space.

Firstly, this study assumes that V is a plane in the n-dimensional inner product space with normal vector $\vec{n}$ and arbitrary point $\vec{r}_0$. The equation of plane V is presented as $\left\langle \vec{n}, \vec{r} - \vec{r}_0 \right\rangle = 0$, in which $\langle \cdot, \cdot \rangle$ is a vector inner product. In an arbitrary vector $\vec{r}_i, i \in 1, 2, \ldots, n$ in an n-dimensional inner product space, the inner product between $\vec{n}$ and $\vec{r}_i$ with virtual vector y (including entries -1 and 1) is expected to be $y_i \left\langle \vec{r}_i, \vec{n} \right\rangle \le a_i < \tilde{a}, i \in 1, 2, \ldots, n$ when $\tilde{a} = \max\{a_i\}$. The optimal $\widehat{a}$ is the objective function, and let $\widehat{a} = \tilde{a} / \| \vec{n} \|$.

$$\begin{aligned} & Min \quad \widehat{a} \\ & S.t., \; y_i \left\langle \vec{r}_i, \vec{n} \right\rangle \le a_i < \widehat{a}; \quad i \in 1, 2, \ldots, n. \end{aligned} \tag{14}$$

Equation (14) can be transformed into the following equation:

$$\begin{aligned} & Max \quad \tfrac{1}{2} \| \vec{n} \|^2 \\ & S.t., \; y_i \left\langle \vec{r}_i, \vec{n} \right\rangle < 1; \quad i \in 1, 2, \ldots, n. \end{aligned} \tag{15}$$

Secondly, a Lagrange multiplier α is introduced to the objective function of (15) to solve the model (Appendix B). The dual Lagrange optimisation is as follows:

$$\begin{aligned} & \underset{\alpha}{Max} \quad \sum_{i=1}^{n} \alpha_i - \sum_{i,j=1}^{n} \alpha_i \alpha_j y_i y_j \vec{r}_i^{\,T} \vec{r}_j \\ & \quad S.t., \; \alpha_i > 0; \quad i \in 1, 2, \ldots, n. \end{aligned} \tag{16}$$

Solution α_i^* can be obtained through sequential minimal optimisation or other iteration algorithms [36]. The inner product space is determined by a normal vector of hyperplane.

$$\vec{n}^* = \sum_{i=1}^{n} \alpha_i * y_i \vec{r}_i$$

Finally, a kernel function is used to deal with the linearly inseparable data by transforming them into a higher dimensional inner space. The Gaussian kernel function is commonly used to map lower dimensional data into a higher dimensional inner product space because the preference relation values obey normal distribution and satisfy normal property [13]. The kernel is $K\left(\vec{r}_i, \vec{r}_j\right) = \exp\left(-\frac{\| \vec{r}_i - \vec{r}_j \|^2}{2\sigma^2}\right)$, and the optimisation model is as follows:

$$\begin{aligned} \underset{\alpha}{Max} \quad & \sum_{i=1}^{n}\alpha_i - \sum_{i,j=1}^{n}\alpha_i\alpha_j y_i y_j K\left(\vec{r}_i, \vec{r}_j\right) \\ S.t., \ & \alpha_i > 0; \quad i \in 1, 2, \ldots, n. \end{aligned} \tag{17}$$

5. Group consensus reaching process

When the number of participants is large, a group consensus is difficult to directly achieve by aggregating each participant's preference. This task needs some consensus reaching approaches to motivate nonconsensus DMs to change their preferences. To deal with this problem, this section designs a three-step consensus reaching process, which is the third step of the proposed framework in Section 2. The first step dynamically generates alternatives based on the subgroups' opinion in a social network. The second step builds an optimisation model to minimise the total compensation cost incurred during the consensus reaching process. The third step generates a ranking of alternatives with the optimal consensus cost.

5.1. Dynamic generation of alternatives

In many LSSNGDM situations, alternatives are dynamically developed along with the changes of the external environment [18]. Negotiations would fail if the alternatives are completely unacceptable by DMs [8]. Approaches have been developed to dynamically generate alternatives, such as medical consultation under mobile and Web 2.0 [28]. In a local social network, alternatives can be provided by a moderator or DMs (Fig. 6). A moderator is trusted by most participants in LSSNGDM, and he/she can propose some initial alternatives after discussing with all involved parties. Meanwhile, DMs can also propose their initial alternatives. These two sets of alternatives are discussed by all participants to generate an initial set of alternatives by eliminating unreasonable alternatives with few supporters and keeping feasible alternatives backed up by the majority of people. Fig. 6 shows the dynamic generation of alternatives for a real-life LSSNGDM problem: an urban resettlement project.

We propose an aggregation operator for different preference forms in an inner product space for each node v in V. The detailed process is shown in Appendix C. The collective opinion of a LSSNGDM can be aggregated from the opinions of different communities by using the optimisation model (Appendix C). The weights are the centrality degree of each community in a local social network calculated using Equation (8).

5.2. Minimum consensus cost

Misclassifications can lead to cost biases and economic losses [44] in the consensus reaching process because they increase the number of participants with nonconsensus preferences in subgroups/classes.

If a consensus cannot be achieved during the generation of the initial alternatives in the first step (Section 5.1), then DMs with nonconsensus opinions are given suggestions to modify their preferences. One approach is to seek feasible strategies with minimum cost paid to nonconsensus DMs, who reject to modify their preferences because the compensation does not satisfy their demands [6,16]. This subsection determines the minimum cost that should be paid to each DM to modify his/her preference [14].

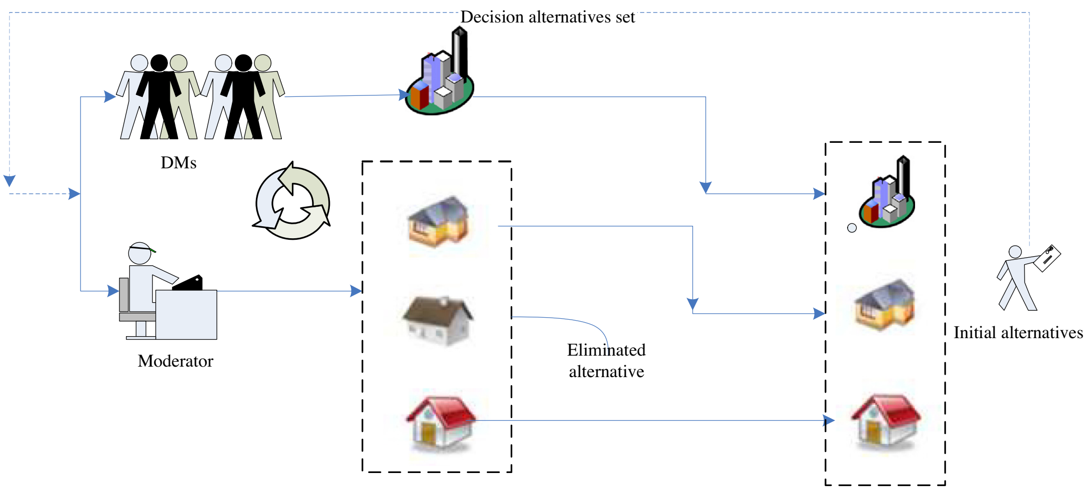

Fig. 6. Dynamic generation of initial alternatives.

Let r_i be a unit compensation cost paid to the i^{th} DM DM_i. $o_C = (o_{c1}, o_{c2}, ..., o_{cn})$ is the order of alternatives based on the group opinion where n is the number of alternatives. $(h_{ii}^v, h_{i2}^v, ..., h_{in}^v)$ represents the order of alternatives by DM_i^v who belongs to the v^{th} subgroup. Then $\sum_{j=1}^{n}\left|h_{ij}^v - o_{cj}\right|$ is the deviation of DM_i s opinion from the group opinion where j is the ranking of the j^{th} alternative. The total optimal cost for reaching a consensus is as follows [16]:

$$\begin{aligned} Min \quad & \sum_N \sum_{i=1}^{m_v} r_i \sum_{j=1}^{n}\left|h_{ij}^v - o_{cj}\right| \\ S.t. \quad & \{o_C \in O \end{aligned} \tag{18}$$

where O is a feasible ranking of alternatives. The solution $(o_{c1}, o_{c2}, ..., o_{cn})^T$ is nonempty and finite and can be easily solved. The impacts of DMs on the collective opinion vary due to their different influences in social networks [12]. The collective opinion is generally closer to communities with larger centralities because they have a greater influence in social networks. Communities with less centralities have less influence on the collective opinion, and the deviations of their group opinions from the collective opinion are larger than communities with higher centralities. Therefore, the moderators must identify the centralities of communities in a social network and set the range of individual preference adjustments according to their influences. A greater degree of preference adjustments is needed for communities with lower centralities to encourage consensus reaching. Meanwhile, less degree of preference adjustments is needed for communities with larger centralities. With regard to a given consensus degree ξ, which is introduced in Appendix D, the total sum of deviations of all DMs in each subgroup is limited by the following deviation control:

$$\xi_v = n^2 s_v^{-1} \xi \tag{19}$$

where ξ_v is a distributed consensus degree calculated using s_v in Equation (8). Each DM in a subgroup can modify his/her preference within a controllable interval. The minimum cost model is established as follows:

$$\begin{aligned} & Min \quad \sum_N \sum_{m_v} r_i \sum_{j=1}^{n}\left|h_{ij}^v - o_{cj}\right|, \\ & S.t. \quad \begin{cases} o_C \in O; (20-1) \\ \frac{1}{Nm_v n^2} s_v \sum_{j=1}^{n}\left|h_{ij}^v - o_{cj}\right| \le \xi (20-2) \\ \left|h_{ij}^v - o_{cj}\right| \le \xi_v; i \in \{1,2,...,m_v\}, v \in \{1,2,...,N\} (20-3) \end{cases} \end{aligned} \tag{20}$$

where O is a feasible ranking of alternatives, in which O is the aggregation of individual preferences. r in Equation (20) indicates the unit compensation cost, which is paid to DMs to encourage them to adjust their preferences to approaching the collective opinion. The two compensation strategies are as follows: the first one is to adopt the same unit compensation cost for each DM. In this case, the total cost is the product of the unit cost and the total opinion deviation, which is the distance between the individual preference and the collective opinion [16]. The other strategy applies different unit compensation costs for different DMs, according to the economic cost of different alternatives. For example, in an urban resettlement project, different unit compensation costs are set according to the economic costs of different demolition plans. Another example is in peer-to-peer online loans. A platform can set different compensation costs to borrowers according to their credit status to reach an agreement between lenders and borrowers [45]. Misclassifying DMs will lead to larger deviations in the objective function. Condition (20–2) means that the total consensus degree should be less than a given ξ. Condition (20–3) represents that each DM should modify his/her preference in terms of his/her centrality defined in Equation (19).

Model (20) can be transformed into a weighted linear integer programming by replacing $\left|h_{ij}^v - o_{cj}\right| = u_{ij}^v + v_{ij}^v$ and $h_{ij}^v - o_{cj} = u_{ij}^v - v_{ij}^v$, where u_{ij}^v and v_{ij}^v are non-negative real numbers [16]. Specifically:

$$\begin{aligned} & \mathrm{Min} \quad \sum_N \sum_{m_v} r_i \sum_{j=1}^{n}\left(u_{ij}^v + v_{ij}^v\right), \\ & S.t. \begin{cases} o_{cj} + u_{ij}^v - v_{ij}^v = h_{ij}^v, i \in \{1,2,...,m_v\}, j \in \{1,2,...,n\} \\ s_v \sum_{j-1}^{n} u_{ij}^v - v_{ij}^v \le \xi N m_v n^2 \\ o_{cj} \le h_{ij}^v + \xi_v, i \in \{1,2,...,m_v\} \\ o_{cj} \le h_{ij}^v - \xi_v, i \in \{1,2,...,m_v\} \\ u_{ij}^v \ge 0, v_{ij}^v \ge 0, \xi_v \ge 0, o_c \in O. \end{cases} \end{aligned} \tag{21}$$

The solution of this linear system is:

$$X = \left(o_{c1}, o_{c2}, ..., o_{cn}; u_{11}^v, u_{12}^v, ..., u_{1n}^v; ..., u_{m1}^v, u_{m2}, ..., u_{mn}^v; v_{11}^v, v_{22}^v, ..., v_{2n}^v, ..., v_{m1}^v, v_{m2}^v, ..., v_{mn}^v\right)^T, v \in \{1,2,...,N\}. \tag{22}$$

Table 2
Consensus reaching algorithm.

Consensus reaching algorithm
Input: initial local social network V; adjacent matrix A; initial consensus degree ξ; heterogeneous preference relations of DMDM_i^v; unit cost r_i.
Output: $X, X', \bar{X}$
Community detection // spectral clustering
Train classification model // inner product SVM
for $v = 1:1:M_v$ **do**
Compute s_v; // weighted eigenvector centrality based on dual network (Equation (8))
end
for $i \in DM$ **do**
Label DM_i to node M_v; // establish the subgroups belonged to each community in the local social network
end
for $i = 1:1:m_v$ **do**
Compute $r_i \sum_{j=1}^{n} (u_{ij}^v + v_{ij}^v)$ // minimum cost for each participant to modify his/her preference
end
return to the selection procedure

However, condition $\left|h_{ij}^v - o_{cj}\right| \le \xi_v$ in (20–3) may lead to no feasible solution if the deviations in different subgroups surpass the preset thresholds. For example, if we need to obtain the optimal value o_c between $h_1 = 1$ and $h_2 = 10$ and restrict $|h_1 - o_c| < 3$ and $|h_2 - o_c| < 3$, then the optimal solution o_c does not exist. In this case, the subgroups are regarded as a non-consensus class, and the cost for consensus reaching in this case is the total deviation computed by$\sum_N \sum_{m_v} r_i \sum_{j=1}^{n} \left|h_{ij}^v - o_{cj}\right|$.

The results of the following matrixes are the ranking of alternatives and the corresponding optimal consensus costs $r_i \sum_{j=1}^{n} (u_{ij}^v + v_{ij}^v)$:

$$\bar{X} = \left[\underbrace{M_v}_{Subgroup} \middle| \underbrace{X'}_{consensus} \right] \tag{23}$$

and

$$X' = \left[\underbrace{o_c^v}_{\text{Ranking}} |M_v \middle| \underbrace{r_i \sum_{j=1}^{n} (u_{ij}^v + v_{ij}^v)}_{\text{Costs}} \right] \tag{24}$$

Equation (24) is the consensus cost of each participant in a subgroup belonged to community v to reach a group consensus.

5.3. Consensus reaching algorithm

The proposed consensus reaching model includes three main procedures: computing weighted eigenvector centrality, classifying preference relations and minimising consensus cost. Table 2 summarises the consensus reaching algorithm.

6. Urban resettlement case study

Urban resettlement has been a popular research topic in the last few years [1,16,25,39]. Urban resettlement projects are typical LSSNGDM problems because the number of stakeholders is large, and residents who are affected by a resettlement project often have tight social connections. Another feature of urban resettlement is the requirement of a high group consensus because it is related to the important assets of most families, and even a few disagreements can fail a project. This situation indicates that minority opinions in urban resettlement projects cannot be ignored.

The example used in this paper is the resettlement of '69 mail box' residence area built in 1950 s in Chengdu, China. The resettlement is a commercial project involving 1861 households, a real estate developer and a government representative who acts as a moderator in the GDM process. The local government stipulated that the resettlement plan must be agreed by all the households to ensure democracy and satisfaction of all the participants, which means that the project cannot proceed if any household does not support the plan. The project was originally initiated in 2007, but it was terminated because the residents and the real estate developer could not reach an agreement. After an initial survey from February 23 to March 5, 2017, a simulated demolition meeting was hold on June 18, 2017, and the results showed that all residences agreed to the demolition. This project was restarted in July 2017 as part of a large government-supported plan 'northern old town improvement'. The goal of the GDM process in this urban resettlement project is to come up with a resettlement plan or a set of plans that consider all residents' opinions and determine rational compensation costs, which are acceptable by the residents and the real estate developer.

The GDM process of this project has four steps. Firstly, the real estate developer proposed alternatives, and the moderator convened 20 representatives, including a neighbourhood committee, resident representatives and a government officer, for discussion. A neighbourhood committee is a self-governing organisation at the grassroots level in urban areas to provide various types of public services to local communities. The committee is elected by local residents every 3 years and consists of five to nine members. In this project, three members were invited to be representatives, and they coordinated with the households during the whole negotiation. Secondly, all the residents have a thorough examination of the alternatives developed in the first step and expressed their preferences. Thirdly, the residents were divided into different subgroups according to their preferences and were provided with the corresponding compensation proposals. Finally, the moderator analysed the consensus and total costs and decided whether this project can proceed to the implementation stage.

This urban resettlement project exhibits some common features of a LSSNGDM: 1) A large number of participants (more than 1000) are involved in the GDM problem, and the residences have complex social relations, which are difficult to fully investigate and display. 2) Given that the deviation of each individual's preference to a collective opinion varies, different consensus costs are needed to persuade individuals to change their preferences to accept a group preference.

6.1. Basic features

An individual preference relation is the preference expressed by a DM for a set of alternatives. This relation is a vector or a matrix with reciprocal or complementary features. In the urban resettlement, the preference of a household is expressed through the ranking of the alternatives, the utility of the alternatives or the pairwise comparison of the alternatives. The preference can be a simple ranking or utility value of the alternatives by directly assessing the importance of the alternatives, which is a permutation function over the set of alternatives X $\mu : X \rightarrow E$ where E is the domain of representation of ranking and utility value of the alternatives. A pairwise comparison matrix is a matrix form of individual preference relationship. The aforementioned matrix compares the importance of the alternatives in pairs, and the scale value constitutes a reciprocal or complementary matrix. In this case, a preference relation on the set of alternatives X is a binary relation $f : X \times X \mapsto D$, where D is the domain of representation of preference degrees provided by the DM. In n alternatives, the preference relation constitutes an $n \times n$matrix, in which entry $f(x_i, x_j)$ is the degree or intensity of the DM's preference of alternative x_i over x_j. The details of preference relations are presented in Appendix A.

All preference relations are vectorised using the cosine similarity relation to visualise their distribution (Appendix A). Then, we build the inner products of all the preference relations by normalising the vectors in Equations (9)–(12). Finally, we reduce the dimensionality of the vectors and represent it in a 2D space using the t-SNE algorithm [26] which is used to observe the distribution of high-dimensional data. The basic idea of the t-SNE is to affine transform the distance between high-dimensional data into conditional probability to express the proximity degree between data. These data are then constructed in a 2D space with a probability distribution as similar as possible to their high-dimensional space. Fig. 7 describes

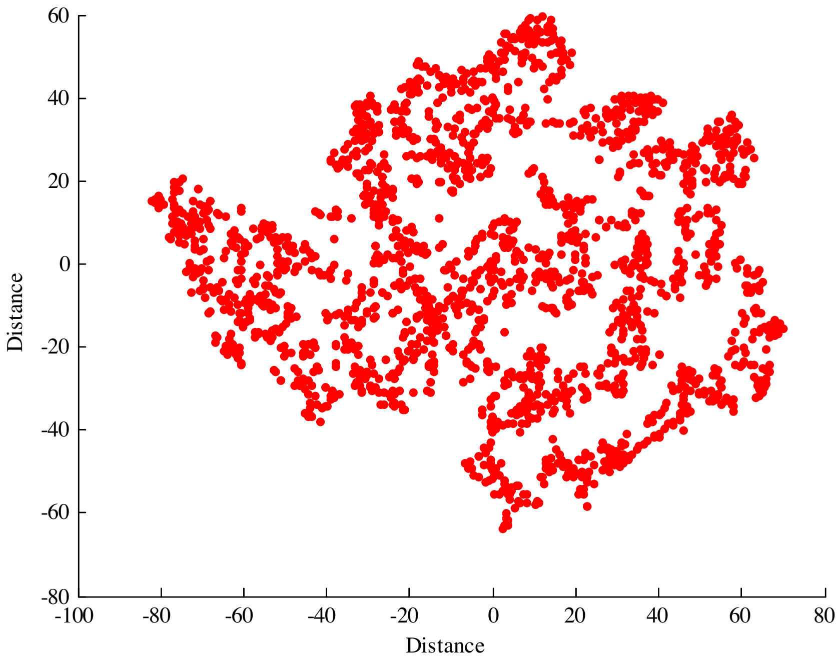

Fig. 7. Distribution of individual preference relations in the urban resettlement project.

the distribution of individual preference relations of the 1861 households. The residents' preference relations exhibit a wide range of distribution because the points are scattered in an interval $[-90, 80]$ which means that the households' opinions are quite different. This situation increases the difficulty of reaching a group consensus.

The participants of this urban resettlement can be divided into three parts (Table 3). A government officer is the moderator whose role is to facilitate the formation of a consensus between the residents and the real estate developer. Residents are the largest part of participants. These individuals want to maximise their compensations for loss of assets and expect better living conditions. The real estate developer expects to make a profit by selling apartments built on the demolition site and hope to pay the minimum resettlement cost to reach a consensus.

A local social network is formed by 20 representatives from the neighbourhood committee, the real estate developer and the residents' representatives. Table 4 summarises the initial alternatives that were formed by taking inputs from the 20 representatives.

Table 5 summarises the four types of preference relations and their usage frequency in the case study. The additive and multiplicative preference relations, which contain richer information than simple ranking, were used more often than the other two types. The additive or multiplicative preference relation can reveal the degree of preference differences for different alternatives. Therefore, the use of preference relations can make negotiation and adjustment of individual opinions more targeted and efficient.

The entire process is implemented in three steps: construct a dual local social network, classify participants into subgroups and reach a consensus with minimum cost. The results of each step are illustrated in the following subsections.

Table 3
Members in the '69 mail box' large-scale GDM.

Participants	Numbers
Moderator	1
Government officer	1
Local social network	20
Neighbourhood committee	3
Real estate developer	1
Residents' representatives	16
Outside of the local social network (other residents)	1842

Remark: a family was treated as a basic unit in the case study. Each family chose a representative to participate in the negotiation.

Table 4
Initial alternatives.

Serial	Alternatives	Notations
a	Resettlement at the original address with purchase of extra construction areas	Replacement at the original address and pay RMB 17,800/m^2 for areas beyond the original construction area
b	Resettlement at the original address	Replacement at the original address with the same size
c	Reform by self-organisation	Select a real estate developer by democratic consultation of all the residences
d	Cash compensation	Compensate RMB 17,000/m^2 for demolished houses and provide the rental fees during the transition period
e	Replacement of new houses at an offsite address	Provide houses of the same size at a new location, which is far away from the original address

Table 5
Summary of preference relations.

Preference formats	Total number	Proportion
Multiplicative	580	31.2%
Additive	740	39.8%
Utility	254	13.6%
Ordering	287	15.4%
Total	1861	100%

Remark: The residents can either choose a simple ranking or make pairwise comparisons. If the residents choose to make pairwise comparisons, the we record their judgments and convert them into appropriate preference relations. The residents can also directly give the utility value of each alternative or simply rank them.

6.2. Main results

This subsection presents the results of applying the proposed framework to the urban resettlement project. We firstly construct a local social network composed of the 20 representatives. If two representatives have a trust relation, which are investigated through interviews and data analysis, then the two nodes representing them have a connection in this network. The preferences of the residences were collected at the simulated demolition meeting on June 18, 2017. All the DMs are divided into small subgroups on the basis of their preference similarities. Then, we establish a local dual network, which is built with trust relations and preference relations. The weights of the subgroups are determined by their centralities. Thereafter, we build an optimisation model to obtain the optimal consensus costs for each subgroup. Lastly, the real estate developer discusses the collective opinion and compensation costs with each household to reach an agreement.

6.2.1. Construct a local social network

In the first step, social relationships between individuals are investigated through data analysis and interviews. Then, we establish trust relations in the local social network, in which their weights and directions are assigned using the above-mentioned data analysis and interviews (Fig. 8). Specifically, the data used in the case study include the basic information (including the household income, educational level and occupation), the trust relations and the preference relations of the 20 representatives. All these data were collected through interviews. We used RMB 60,000, the median of the 19 households' income (except the representative of the real estate developer) as the boundaries, to divide the households to obtain the trust relationships between different types of households. Then, we combined the partial trust relationships obtained from the interviews and constructed a local trust network. Fig. 8 illustrates the network graph of the local trust relation network.

6.2.2. Subgroup classification

All the DMs are divided into small subgroups based on the following two steps.

Step 1: Community detection. This step divides the local social network into communities based on the connections between nodes. The local social network in this example has 20 nodes, including the neighbourhood committee, the real estate developer and the representatives of households.

The preference similarity matrix and the degree matrix are deconstructed into a Laplace matrix using Equation (13). The Laplace matrix serves as a project vector of each node (Equation (13)) and is used to detect communities. The Laplace matrix includes the preference similarities and trust relations. Thus, DMs with trust relations and different preferences may also be assigned in one community. The number of clusters is five, which is determined by the number of initial alternatives (Table 4). The detailed process is discussed in Section 6.2.3. The clustering results are as follows:

$$\left(\begin{array}{c|ccccc|c}
\overbrace{\text{Serial number}} & \multicolumn{5}{c|}{\overbrace{\text{Indiviual Preferences(priority vector)}}} & \overbrace{\text{Cluster number}} \\
1 & 0.0588 & 0.2984 & 0.0348 & 0.4084 & 0.1995 & 4 \\
2 & 0.4137 & 0.1110 & 0.2096 & 0.1978 & 0.0679 & 2 \\
3 & 0.2883 & 0.0535 & 0.4595 & 0.1696 & 0.0291 & 5 \\
4 & 0.4315 & 0.1348 & 0.1054 & 0.1239 & 0.2044 & 2 \\
5 & 0.0295 & 0.3635 & 0.1401 & 0.2079 & 0.2590 & 1 \\
6 & 0.0416 & 0.4648 & 0.0814 & 0.0895 & 0.3227 & 1 \\
7 & 0.3020 & 0.0561 & 0.3934 & 0.0324 & 0.2161 & 5 \\
8 & 0.1602 & 0.1968 & 0.3008 & 0.2364 & 0.1057 & 5 \\
9 & 0.3942 & 0.1305 & 0.0315 & 0.3942 & 0.0496 & 3 \\
10 & 0.0745 & 0.3789 & 0.0553 & 0.1006 & 0.3907 & 1 \\
11 & 0.3871 & 0.1414 & 0.2099 & 0.0501 & 0.2114 & 2 \\
12 & 0.0431 & 0.0638 & 0.3861 & 0.4184 & 0.0886 & 5 \\
13 & 0.3157 & 0.0542 & 0.0737 & 0.4266 & 0.1298 & 3 \\
14 & 0.0531 & 0.1848 & 0.2672 & 0.0237 & 0.4713 & 1 \\
15 & 0.1328 & 0.0540 & 0.1513 & 0.5762 & 0.0856 & 3 \\
16 & 0.3763 & 0.2958 & 0.0705 & 0.1745 & 0.0830 & 2 \\
17 & 0.1457 & 0.2877 & 0.0530 & 0.2927 & 0.2209 & 4 \\
18 & 0.5901 & 0.1092 & 0.1988 & 0.0517 & 0.0501 & 2 \\
19 & 0.0829 & 0.0682 & 0.4686 & 0.1479 & 0.2324 & 5 \\
20 & 0.0688 & 0.5453 & 0.0481 & 0.1130 & 0.2248 & 1
\end{array}\right) \tag{25}$$

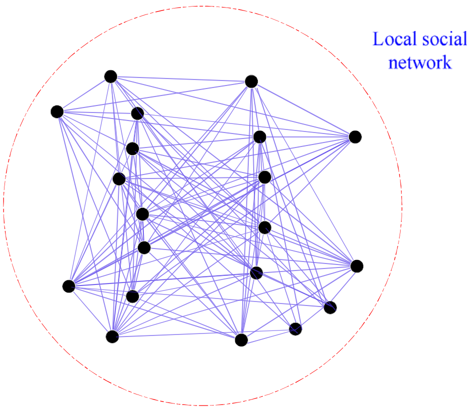

Fig. 8. Constructed local social network for the '69 mail box' project.

Table 6
Communities in the local social network.

Labels	Representative opinion	Ranking/order					Number of nodes
		a	b	c	d	e	
RS	Cash compensation with resettlement at the original address without extra expenditure	0.2020 4	0.5652 2	0.0857 5	0.6705 1	0.4068 3	5
BU	Resettlement (replace or buy) at the original address	0.8057 1	0.2975 2	0.2926 3	0.2263 5	0.2347 4	5
CB	Buy new houses after cash compensation or through self-financing	0.4899 2	0.1373 5	0.1442 4	0.7954 1	0.1529 3	3
CR	Replacement of house property (original address or offsite address)	0.0935 5	0.6811 1	0.2111 3	0.1933 4	0.5895 2	2
RF	Self-organisation reform	0.3245 3	0.1694 5	0.7349 1	0.3689 2	0.2487 4	5

Remark: the real estate developer is included in the community *CB*. The group opinion is integrated with weighted centrality degree.

Table 6 shows the five communities in the local social network. The ranking/order column indicates the collective opinion of each community about the five alternatives (Table 4), and the order is the same ranking result using Arabic numerals. The right most column is the number of nodes in different communities. Appendix C provides the calculation details about how the individual preferences are aggregated to form the ranking.

Each community in the local social network has its typical opinion, which is represented by the label column. *RS* represents a preference of cash compensation to buy a new house at the original address of the same size. Households in community *BU* want larger new houses at the original address, and they are willing to pay for the areas that exceed the construction areas before the demolition. *CB* community hopes to obtain enough cash compensation that can be used to buy a new house at the original address with extra construction areas. *CR* community wants a replacement of the same size at the original or offsite addresses. *RF* wants to change the real estate developer and find a new one who can pay higher compensations.

Although the labels in Table 6 and the alternatives in Table 4 have similarities, they are different. The labels in Table 6 are obtained by integrating DMs' preferences of the alternatives in each community. We analyse the collective opinion and the top-ranking alternatives of each community to determine the preference characteristics of that community and assign labels to communities using their main characteristics. For example, the top three ranking alternatives in the first community of Table 6 are {d, b, e}. A common characteristic of these three alternatives is their reluctance to pay extra cost.

Step 2: Divide DMs into subgroups. The five communities identified in step 1 are used to train the inner product SVM. This step classifies the outside nodes, which are the majority of the participants in this project, into the five subgroups.

In this step, each community is a class: {*RS, BU, CB, CR, RF*}, using the labels from Table 6. Individual preferences are transformed into an inner product space (Appendix A). The classification is implemented using Equation (17) to classify the heterogeneous preferences in LSSNGDM.

Table 7 presents the classification results. The largest subgroup is *RS*, and the smallest one is *BU*. Although *CB* and *CR* together account for only 25% of the total number of representatives in the local social network (Table 6), they represent 37% of the participants in the whole social network.

The classification results are used to determine the structure of the dual network and evaluate the weighted eigenvector centrality of each community. The centrality of one node is the degree of the connections from this node to other nodes in a

Table 7
Classification results.

Subgroups	Number of representatives	Size of subgroups	Total	Centrality
RS	5	522	527	1.1404
BU	5	287	292	1.1112
CB	3	374	377	0.6249
CR	2	319	321	0.6259
RF	5	340	345	1.2021
Total:5	20	1842	1862	

social network and indicates the importance or influence of the node. The centrality has two roles in the proposed approach: 1) different subgroups can adjust their preferences to various degrees depending on their centralities; 2) centralities of communities can be used as weights of communities to aggregate their opinions to obtain a collective opinion.

6.2.3. Weights of subgroups

The weights of subgroups, which are the weighted centralities of nodes in subgroups, are calculated by combining the trust relations and the preference relations.

In the first step, the trust relations amongst the DMs are regarded as a base layer. Then, an adjacent matrix of the different subgroups in this layer is built. The adjacent matrix $A' = (a_{ij})$of subgroups is derived from the communities in the local social network, whose transformation principle is as follows:

$$e_{ij} = \sum_{\substack{s \in communityi \\ s \in communityj}} ce_{st} \tag{26}$$

where ce_{st} is the number of linked edges from the s^{th} node to the t^{th} node, and e_{ij} is the number of linked edges in communities i and j. The following adjacent matrix reflects the various connections amongst different communities. For example, $e_{12} = 14$ and $e_{14} = 5$ indicate that more links exist between community *RS* and *BU* ($e_{12} = 14$) than between *RS* and *CR* ($e_{14} = 5$).

$$A = (e_{ij}) = \begin{array}{c} \\ \\ RS \\ BU \\ CB \\ CR \\ RF \end{array} \left(\begin{array}{ccccc} \multicolumn{5}{c}{Communities} \\ \multicolumn{5}{c}{\overbrace{12345}} \\ 0 & 14 & 7 & 5 & 11 \\ 14 & 0 & 6 & 5 & 11 \\ 10 & 4 & 0 & 2 & 4 \\ 3 & 6 & 4 & 0 & 7 \\ 12 & 15 & 9 & 3 & 0 \end{array} \right) \tag{27}$$

This matrix can then be combined with a participant's coefficient vector C using Equation (1), which is a normalised vector of the total column in Table 7.

$$C = (0.2830 \quad 0.1568 \quad 0.2025 \quad 0.1724 \quad 0.1853)^T \tag{28}$$

The coefficient of the base layer can be obtained as follows:

$$A' = AC = (\ 0.2093 \quad 0.2596 \quad 0.1460 \quad 0.1252 \quad 0.2599)^T \tag{29}$$

In the second step, we construct a preference similarity matrix for different communities. The entries in this matrix are cosine similarity relations of representative opinions, which are the third column in Table 6. For example, the preference similarity of the community *RS* and *BU* is 0.5699. Specifically:

$$D' = (d_{ij}) = \begin{pmatrix} 1.0000 & 0.5699 & 0.4503 & 0.8405 & 0.5850 \\ 0.5699 & 1.0000 & 0.7437 & 0.6312 & 0.7428 \\ 0.4503 & 0.7437 & 1.0000 & 0.8181 & 0.6860 \\ 0.8405 & 0.6312 & 0.8181 & 1.0000 & 0.6189 \\ 0.5850 & 0.7428 & 0.6860 & 0.6189 & 1.0000 \end{pmatrix} \tag{30}$$

and the coefficient of the preference similarity is computed by C:

$$D = D'C = (d_{ij}) = (0.1964 \quad 0.1960 \quad 0.1959 \quad 0.2164 \quad 0.1954)^T \tag{31}$$

The combination of two vectors $S = [A'|D]$ can be decomposed by singular value decomposition using Equation (13):

$$\begin{pmatrix} \theta_1 & . \\ . & \theta_2 \end{pmatrix} = \begin{pmatrix} 1.3981 & . \\ . & 0.2131 \end{pmatrix} \tag{32}$$

Fig. 9 is the dual network, in which nodes with the same colour belong to one community:

Finally, we calculate the weights of the subgroups using Equation (8), which is the weighted eigenvector centrality of each community (the last column in Table 7):

$$\vec{s} = (s_v) = \begin{array}{c} RS \\ BU \\ CB \\ CR \\ RF \end{array} \begin{pmatrix} 1.1404, \\ 1.1112, \\ 0.6249, \\ 0.6259, \\ 1.2021. \end{pmatrix} \tag{33}$$

Although *BU* has the smallest number of nodes, its weight is not the least. The highest centrality is *RF*, which is the third largest subgroup. The preferences of subgroups *RF*, *RS* and *BU* have more influence in the social network than those of *CB* and *CR*.

6.2.4. Group consensus process with minimum cost

A unanimous consensus is hard to reach due to certain factors, such as conflicting preferences, asymmetric information and noncooperative DMs. Instead of a complete agreement, consensus building with compensation cost [6,1750] tries to reach a certain degree of group consensus by compensating participants. This mechanism is a realistic approach in many real-world GDM problems. In urban resettlement, most residents are willing to change their preferences if they receive adequate compensation. The challenge is how to determine the optimal compensation cost that can stimulate residents to reach a consensus under the total budget.

This subsection presents our approach of consensus reaching with minimum cost for the urban resettlement project. Firstly, we set a preference deviation threshold ξ_V using Equation (18). The threshold controls an interval $\left|h_{ij}^v - o_{cj}\right| \leq \xi_v; i \in \{1, 2, \ldots, m_v\}, v \in \{1, 2, \ldots, N\}$ in optimisation model (21), and this interval determines the range within which the DMs' preferences can be modified. A subgroup with a higher ξ_V is allowed to make less change of their preferences. This condition is similar to the real-world GDM situations. Important DMs are less likely to change their preferences, and they have a greater influence on the collective opinions. Secondly, the solution $(o_{c1}, o_{c2}, \ldots, o_{cn})^T$ of Equation (22) is the collective opinion, and the consensus cost compensated to the DMs is computed by $r_i \sum_{j=1}^{n} \left(u_{ij}^v + v_{ij}^v\right)$.

We set the total preference deviation threshold $\xi \leq 0.32$ on the basis of management experience, which controls the total deviation of individual preferences from the group collective opinion. We choose this threshold because more than 70% of the residents in the '69 mail box' resettlement project agreed on the demolition, which initiated the simulation demolition meeting. In this case, the value 0.32 of ξ is calculated according to Equation (D-3) in Appendix D. The average deviation of different subgroups ξ_v is determined by the weighted centrality degrees. The optimal compensation cost for each subgroup is

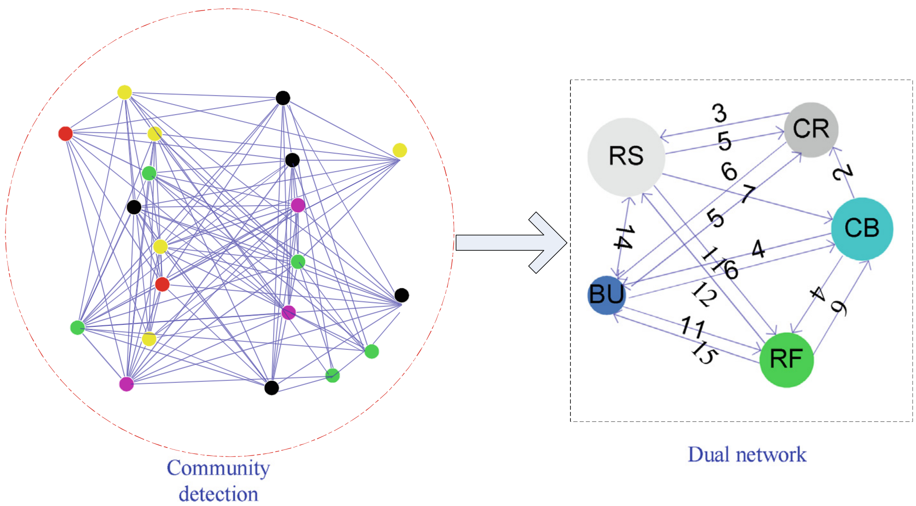

Fig. 9. Local and dual networks.

Table 8
Optimal results.

Labels	Costs	Centrality	ξ_v		
RS	1898	1.1404	7.0149		
BU	2130	1.1112	7.1997		
CB	3630	0.6249	12.8011		
CR	3970	0.6259	12.7811		
RF	2920	1.2021	6.6552		
Collective group opinion	**b**	**a**	**c**	**e**	**d**
				Total cost	14,548

Remark: In our consensus reaching framework, the collective group opinion is presented as an order of alternatives, rather than their weights, because the goal of this selection process is to reach consensus while minimising consensus cost. However, this mechanism cannot show the gap between different alternatives. In Equation (18), r_i is a unit compensation cost, and it is a dimensionless unit. Mathematically, this variable is the coefficient of deviation between individual preference and the collective opinion. Thus, the 2th column in Table 8 (and all the costs in Section 6) has no unit.

calculated using optimisation model (21). The collective opinion is ranked as $b \succ a \succ c \succ e \succ d$, and the explanations of the alternatives can be found in Table 4.

Table 8 shows the total cost and the preference deviation ξ_v of each subgroup. Variable ξ_v monotonically decreases when the centrality degree increases, which means the higher ξ_V of a subgroup, the smaller its preference needs to be modified. Given that the cost here is the compensation paid to the DMs to change their preferences, a subgroup with a higher centrality degree will receive less compensation due to their less preference modifications. Meanwhile, subgroups *CB* and *CR* receive more compensation because they have lower centrality degrees, and their original preferences are far from the collective opinion of the group.

Remark:. Although providing the same plan and compensation to every household sounds fair and simple, it doesn't work well in real-life urban resettlement projects. Households have diverse situations and want different alternatives. For example, in the '69 mail box' resettlement project, residents have different expectations for compensation not only because they have varying preferences but also because their apartments have diverse sizes, types, floors and orientations. In addition, the local government stipulates that the resettlement plan must be agreed by all the households, which means that the project cannot proceed if any household does not support the plan. All these factors make it difficult to for all the residents to agree with the same compensation plan. Existing studies support that dividing residents into subgroups can improve management efficiency, increase residents' satisfaction and achieve final consensus [25]. In real-life urban resettlement projects, residents were divided into different groups and treated according to their personal identities and residence status and were compensated differently [30].

6.2.5. Alternative selection process for each household

The collective opinion and the compensation cost obtained from the previous steps are provided to each household. The real estate developer needs to negotiate with different subgroups about their compensation. The cost–consensus matrix is as follows:

$$\left[\begin{array}{c|ccccc|c|c}
\overbrace{RS}^{subgroups} & \multicolumn{5}{c|}{\overbrace{d \quad b \quad e \quad a \quad c}^{subgroup \quad opinions}} & \overbrace{b}^{Collective \quad opinion} & \overbrace{1898}^{compensation} \\
BU & a & b & c & e & d & a & 2130 \\
CB & d & a & e & c & b & c & 3630 \\
CR & b & e & c & d & a & e & 3970 \\
\underbrace{RF}_{community} & \multicolumn{5}{c|}{\underbrace{c \quad d \quad a \quad e \quad b}_{ranking}} & \underbrace{d}_{ranking} & \underbrace{2920}_{costs}
\end{array}\right] \tag{34}$$

The first ranked alternatives in subgroups *CB* and *CR* are the two most expensive alternatives (3630 and 3970). These alternatives represent 44% of the participants, but they account for 52.1% of the consensus costs.

6.3. Comparison analysis

6.3.1. Comparison with methods used in reality

We compare our approach with the current management practices in the '69 mail box' resettlement project in terms of the total consensus cost to verify the effectiveness of the proposed approach. In the actual urban resettlement project, the real estate developer and residences provide their alternatives (Section 5.1) and expect to reach a consensus plan after multiple rounds of negotiations. The compensation plan is guided by an officer from the local government, who averages the preferences of all households and suggests that the real estate developer will equally compensate every household under the budget to ensure fairness and reach a consensus.

Table 9
Total consensus cost comparison.

Methods	Collective group opinion	Total consensus costs	Cost reduction (%)
Average of all the households' preferences	$c \succ d \succ a \succ b \succ e$	14,934	2.34
Real estate developer's plan	$e \succ d \succ a \succ b \succ c$	15,070	3.22
Our approach	$b \succ a \succ c \succ e \succ d$	**14,585**	–

The first row in Table 9 computes the total consensus cost using the average preference of all the households. Given that the collective opinions using the preference means of each subgroup (Table 6) are {0.1981, 0.1972, 0.2084, 0.2022, 0.1941}, the order is $c \succ d \succ a \succ b \succ e$, and the total consensus cost is 14,934. In this project, the initial preference for the alternatives provided by the real estate developer is $e \succ d \succ a \succ b \succ c$, which means that they want to obtain ownership of the land to be demolished through relocating residents at off-site address (e) or cash compensation (d). The second row shows the total consensus cost (15,070) if the real estate developer's plan $e \succ d \succ a \succ b \succ c$ is adopted. The third row presents the results using our approach. The collective group opinion is $b \succ a \succ c \succ e \succ d$, and the total consensus cost is 14,585. Our approach can reduce the cost by 2.34% and 3.22% compared with the two approaches currently used. In this project, the unit compensation cost r_i is RMB 112,806. The economic costs of the three methods, namely, average of all the households' preferences, the real estate developer's plan and our approach, are RMB 168.465, RMB 169.999 and RMB 164.528 million, respectively. Our method reduced by approximately RMB 54.71 million (USD 7.85 million) compared with the business strategy launched by the real estate developer.

A common demolition strategy (the first row in Table 9) averages all residents' preferences, and the alternative ranked the first in this strategy is 'reform by self-organisation' (c). Given that this alternative needs to use a large part of the demolition land to build resettlement departments, it reduces the real estate developer's profit margin. The second common demolition strategy (the second row in Table 9) adopts the solution provided by the real estate developer, who is more interested in developing new commercial departments by fully obtaining the ownership of the original land. Urban resettlement projects often involve lands that are located at the centre of the city and have high commercial value. Thus, the top two alternatives of this strategy are 'Replacement of new houses at an offsite address' (e) and 'Cash compensation' (d), which require a larger cost to compensate the residents to resettle at off-site addresses. The demolition costs of these two strategies are higher than that of our approach.

The diverse results of the three strategies are because of the different rankings result in varying deviation degrees between the collective opinion and the individual preferences. The ranking of different alternatives needs to be optimised to minimise the deviations between the collective opinion and the individual preferences because the distribution of individual preferences greatly varies (Fig. 7). The real estate developer's plan focuses on its own interests and deviates from the wishes of most residents; thus, its cost is higher. Averaging all preference of DMs does not consider the distribution characteristics of individual preferences; thus, it is difficult to reduce costs. The proposed approach can achieve lower compensation costs than the previous two strategies because our framework optimises the demolition cost through the classification of preference similarities and trust relations amongst the residents.

6.4. Comparison with existing consensus reaching methods in social network GDM

Major consensus reaching methods for social network GDM include minimum cost consensus method [7,45] feedback adjustment mechanism [34,38] punishment mechanism [10] and TOPSIS selection [41]. The minimum cost model encourages DMs to adjust their preferences by compensation, whilst the feedback adjustment mechanism provides ranges of feedback preference modification information to DMs for references. These methods adopt different approaches when they deal with varying numbers of DMs. A feedback adjustment mechanism is adopted for less than 25 DMs to ensure consensus reaching through weight adjustments. The minimum cost model is applied to numerical preference relation and less than 20 DMs. A penalty mechanism is often used for larger than 50 DMs to adjust the weights of DMs to achieve rapid consensus convergence. Table 10 shows that the time complexity of various consensus reaching methods is basically at a polynomial level. The feedback mechanism has the lowest time complexity, belonging to the $O(n^2)$ level, because the mechanism does not use optimisation methods. The complexity of other methods is at the $O(n^3)$ level. Therefore, no significant difference exists in the time complexity of these methods. Our method includes three parts: community detection, support vector machine classification and cost optimisation. The complexities for these three parts are $O(n^3 + 2n)$,$O\left(Mn^2 + M^3 + M\right)$ and$O(2n^3 + 3n^2 + 2n)$, respectively. M is the number of DMs, and n is the number of the alternatives. The complexity of the proposed method is approximately $O(3n^3)$.

6.4.1. Comparison with existing subgroup partition methods in social network GDM

The problems studied by existing subgroup partition methods in social network GDM are different in the number of DMs, the form of preference relations and the structure of social networks. Table 11 compares the proposed subgroup partition approach with the existing methods from three aspects:

Table 10
Comparisons of the complexity and characteristic of different methods.

Methods	Preference formats	Social network relation		Large-scale GDM		Consensus reaching process			Calculation complexity
		Trust relation	Preference similarity (distance)	Is it a large group?	Number of DMs	Subgroup partition	Weight allocation	Adjustment strategy	
Minimum cost soft consensus model [45]	Crisp values	✗	✗	No	5	✗	✗	Minimum cost	$O(2n^3+3n^2+2n)$
Fuzzy TOPSIS model [41]	Interval type-2 fuzzy sets	✗	✓	Yes	50	✓	✓	TOPSIS-based selection	$O(n^3+2n^2)$
Similarity–confidence–consistency model [34]	Intuitionistic fuzzy preferences	✗	✓	Yes	25	✓	✓	Feedback mechanism	$O(4n^2)$
Trust propagation model [38]	Interval-valued fuzzy reciprocal preference	✓	✗	No	6	✗	✓	Feedback mechanism	$O(3n^2)$
Social network DeGroot Model [10]	Individual opinion	✓	✗	Yes	200	✗	✓	Penalty mechanism	$O(3n^3+2n)$
Our methods	Heterogenous preferences	✓	✓	Yes	1861	✓	✗	Minimum cost	$O\begin{pmatrix}3n^3+(3+M)n^2\\+4n+M^3+M\end{pmatrix}$

Remark: [34] and [38] need a feedback preference adjustment to achieve consensus convergence.

Table 11
Comparisons of different subgroup partition methods.

Methods	Partition algorithms	Is it applicable to sparse trust relationship?	Can it handle heterogeneous preference information?	Is it applicable to a large-scale GDM problem?
Changeable cluster model [42]	Clustering	No	No	Yes
Two-stage social trust network partition model [40]	Aforementioned shortest path method	No	No	Yes
Subnetwork split model [41]	Louvain method	No	No	Yes
Opinion dynamics model [4]	Opinion dynamic	No	Yes	No
Our methods	Classification	Yes	Yes	Yes

1) Is it applicable to sparse trust relationship? Our approach uses the trust relation and preference similarity in the consensus reaching process compared with the approaches of building social networks through a single network relationship [12,35,38,41]. The complementary nature of the trust relation and preference relation can better reflect the relationship between DMs in social networks.
2) Can it handle heterogeneous preference information? In contrast with the current subgroup partition methods [4,40,41,42] the proposed approach uses a support vector machine based on vector space to form subgroups. This algorithm can not only be used for the heterogenous preferences but also effectively classify decision makers when the trust relationship is partially missing in a large-scale decision-making problem.
3) Is it applicable to a large-scale GDM problem? In contrast with the traditional feedback-adjusted social network GDM consensus framework [34,38] the minimum cost model proposed in this study optimises compensation costs on the basis of the subgroup partition, which makes it not only suitable to obtain the minimum consensus compensation cost but also appropriate for large-scale social network problems under heterogeneous preference information.

6.5. Discussions

Although this study aims to improve the efficiency of consensus reaching for large-scale urban resettlement projects, it can be applied to other real-life GDM problems with the following features:

1) In many real-life large-scale GDM situations, the complete trust relations amongst DMs are difficult to investigate. The proposed approach can be used to facilitate consensus reaching in this case through the topological structure of the local social network and preference classification.

2) When subgroups within a social network have quite different preferences, providing differentiated compensations to different subgroups is more effective and cost-saving. In the actual urban resettlement projects, compensating differently for households with various preferences is often necessary.

3) When the interests of participants are closely related to the decision results, such as urban resettlement (another example is Appendix E), the proposed approach can be used to optimise compensation cost in the consensus reaching process.

7. Conclusions

Urban resettlement projects involve a large number of DMs, who have tight social connections, and these connections have an important impact on the consensus reaching process. LSSNGDM problems are challenging due to the size of DMs (more than 1000 persons) and incomplete social relations. Issues, such as how to determine the structure of social networks, how to classify DMs based on their preferences and how to optimise the consensus costs, are still unsolved.

We proposed a consensus reaching framework for LSSNGDM to tackle these questions. This framework firstly divided large-scale social network into a local network and external suspension nodes. A small portion of the DMs are in the local network with completely known trust relations and preference similarity relations, which are collected through data analysis and interviews. Secondly, we combined the two social relations and identified communities in the local network. Thirdly, we assigned labels to the communities for subgroup divisions. We divided all DMs into subgroups by using the proposed inner product space SVM and determined the consensus compensation mechanism. Finally, we developed a minimum consensus cost optimisation model to guide the consensus reaching process on the basis of the centralities of subgroups. A real-life example of urban resettlement project in China was used to show how the framework works. The results demonstrated that the proposed consensus reaching framework can improve the efficiency and reduce the total cost of consensus reaching in LSSNGDM.

One of our future research directions is to develop a large-scale decision support system in Mobile system and web2.0 to assist real-world LSSNGDM-related management problems.

CRediT authorship contribution statement

Xiangrui Chao: Conceptualization, Methodology, Software, Formal analysis, Writing - original draft. **Gang Kou:** Conceptualization, Investigation. **Yi Peng:** Conceptualization, Visualization, Writing - original draft, Writing - review & editing. **Enrique Herrera-Viedma:** Writing - review & editing. **Francisco Herrera:** Writing - review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors thank the editor and the anonymous referees for their valuable comments and insightful recommendations. This work was supported in part by grants from the National Natural Science Foundation of China (#71874023, #71771037, #71971042, #71910107002, and #71725001) and supported by the Spanish State Research Agency under Project PID2019-103880RB-I00/AEI/10.13039/501100011033.

Appendix A. This appendix introduces four most frequently used preference representation forms and investigates the similarity relation between the preferences and their derived weight vectors. Lastly, the inner product of different preference relations is established.

A preference relation is a function $X \times X$ (alternative set $X = \{x_1, x_2, \dots, x_n\}$) to a preference set D, which is measured by a crisp number or fuzzy value indicating the relative important degree of two alternatives provided by DMs.

1) Utility value. The preference set D is a utility vector of alternatives, and the values in the vector are in the interval $[0, 1]$. The larger the utility value, the more important the alternative is. For example, the utility value {0.6, 0.9, 0.7, 0.4} of alternatives $\{x_1, x_2, x_3, x_4\}$ means that their ranking based on weights is $x_2 \succ x_3 \succ x_1 \succ x_4$.

2) Preference ordering. The preference set D is a ranking of the alternatives in terms of their relative important degrees. For example, $\{3, 1, 2, 4\}$ means that the ordering of the alternatives is $x_2 \succ x_3 \succ x_1 \succ x_4$.

3) Multiplicative preference relation. The preference set D is a preference matrix $(a_{ij})_{n\times n}$. $a_{ij} \in [1/9, 9]$ represents a comparison preference of two alternatives with a relation $a_{ij}a_{ji} = 1$. $a_{ij} = 9$ indicates that the alternative x_i is the most preferred, and x_j is the least preferred when they are compared with each other. Meanwhile, $a_{ij} = 1/9$ means the opposite situation. When $a_{ij} = 1$, the two alternatives are the same for DMs.

4) Additive preference relation. The preference set D is a preference matrix $(a_{ij})_{n\times n}$. $a_{ij} \in [0, 1]$ represents a comparison preference of two alternatives with a relation$a_{ij} + a_{ji} = 1$. $a_{ij} > 0.5$ means that a DM prefers the alternative x_i to x_j. When $a_{ij} = 0.5$, the two alternatives are the same for DMs.

Now, we can illustrate the cosine similarity relation of different preference forms:

1) Utility value: This study assumes that $\overrightarrow{u} = (u_1, u_2, ..., u_n)$ is a utility value for a given alternative set, and $\omega = (\omega_{11}, \omega_2, \ldots, \omega_n)$ is the derived weight vector. The following relation holds when the utility value is consistent with the derived weight vector:

$$S\left(\overrightarrow{u}_j, \omega\right) = \frac{\sum_{i=1}^{n} \frac{u_i \omega_i}{u_j}}{\sqrt{\sum_{i=1}^{n} \left(\frac{u_i}{u_j}\right)^2} \sqrt{\sum_{i=1}^{n} \omega_i^2}} = \frac{\sum_{i=1}^{n} \frac{\omega_i}{\omega_j} \omega_i}{\sqrt{\sum_{i=1}^{n} \left(\frac{\omega_i}{\omega_j}\right)^2} \sqrt{\sum_{i=1}^{n} \omega_i^2}} = 1, \tag{A-1}$$

where $\overrightarrow{u}_j = (\frac{u_1}{u_j}, \frac{u_2}{u_j}, ..., \frac{u_n}{u_j})^T, j = 1, 2, \ldots, n$.

2) Preference ordering. Let $\overrightarrow{o}_j = (o_1, o_2, \ldots, o_n)$ be a preference ordering to evaluate a given alternative set, and $\omega = (\omega_{11}, \omega_2, \ldots, \omega_n)$ is a derived weight vector. The similarity relation holds if the preference ordering is consistent with the derived weight vector:

$$S\left(\overrightarrow{o}_j, w\right) = \frac{\sum_{i=1}^{n} \frac{(n-o_i)\omega_i}{n-o_j}}{\sqrt{\sum_{i=1}^{n} \left(\frac{n-o_i}{n-o_j}\right)^2} \sqrt{\sum_{i=1}^{n} \omega_i^2}} = \frac{\sum_{i=1}^{n} \frac{\omega}{\omega_j} \omega_i}{\sqrt{\sum_{i=1}^{n} \left(\frac{\omega_i}{\omega_j}\right)^2 \sum_{i=1}^{n} \omega_i^2}} = 1, \tag{A-2}$$

where $\overrightarrow{o}_j = (\frac{n-o_1}{n-o_j}, \frac{n-o_2}{n-o_j}, ..., \frac{n-o_n}{n-o_j})^T, j = 1, 2, \ldots, n$.

3) Multiplicative preference relation. In multiplicative preference relation $(a_{ij})_{n\times n}$, $\omega = (\omega_{11}, \omega_2, \ldots, \omega_n)$ is a derived weight vector. The following similarity relation can be hold under the prefect consistency condition:

$$S\left(\overrightarrow{a}_j, \omega\right) = \frac{\sum_{i=1}^{n} a_{ij}\omega_i}{\sqrt{\sum_{i=1}^{n} a_{ij}^2} \sqrt{\sum_{i=1}^{n} \omega_i^2}} = \frac{\sum_{i=1}^{n} \frac{\omega_i}{\omega_j} \omega_i}{\sqrt{\sum_{i=1}^{n} (\frac{\omega_i}{\omega_j})^2} \sqrt{\sum_{i=1}^{n} \omega_i^2}} = 1, \tag{A-3}$$

where $\overrightarrow{a}_j$ is the column vector of multiplicative preference relation. The above equation is discussed in Kou and Lin [21].

4) Additive preference relation. Let $(b_{ij})_{n\times n}$be an additive preference relation, and$(p_{ij})_{n\times n}$ is a transformation matrix of $(b_{ij})_{n\times n}$, where$p_{ij} = b_{ij}/(1 - b_{ij})$. The following relation can be induced if the perfect consistency condition holds:

$$S\left(\overrightarrow{p}_j, \omega\right) = \frac{\sum_i \frac{b_{ij}}{1-b_{ij}} \omega_i}{\sqrt{\sum_i \left(\frac{b_{ij}}{1-b_{ij}}\right)^2} \cdot \sqrt{\sum_i \omega_i^2}} = \frac{\sum_i \frac{\omega_i}{\omega_j} \omega_i}{\sqrt{\sum_i \left(\frac{\omega_i}{\omega_j}\right)^2} \cdot \sqrt{\sum_i \omega_i^2}} = \frac{\sum_i \omega_i^2}{\sqrt{\sum_i \omega_i^2} \cdot \sqrt{\sum_i \omega_i^2}} = 1, \tag{A-4}$$

where $\overrightarrow{p}_j$ is a column vector of $(p_{ij})_{n\times n}$.

Based on the similarity relation between the preference relation and the derived weight vector, the preference relation can be transformed into following unitised matrix.

Let $u_{ij} = u_i/u_j$; then, $(u_i/u_j)_{n\times n}, i,j = 1, 2, \ldots, n$. The column vector is similar to priority vector ω.

$$\bar{U} = \left(\bar{u}_{ij}\right)_{n\times n} = \begin{pmatrix} \frac{u_{11}}{\sqrt{\sum_{i=1}^{n}(u_{i1})^2}} & \frac{u_{12}}{\sqrt{\sum_{i=1}^{n}(u_{i2})^2}} & \cdots & \frac{u_{1n}}{\sqrt{\sum_{i=1}^{n}(u_{in})^2}} \\ \frac{u_{21}}{\sqrt{\sum_{i=1}^{n}(u_{i1})^2}} & \frac{u_{22}}{\sqrt{\sum_{i=1}^{n}(u_{i2})^2}} & \cdots & \frac{u_{2n}}{\sqrt{\sum_{i=1}^{n}(u_{in})^2}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{u_{n1}}{\sqrt{\sum_{i=1}^{n}(u_{i1})^2}} & \frac{u_{n2}}{\sqrt{\sum_{i=1}^{n}(u_{i2})^2}} & \cdots & \frac{u_{nn}}{\sqrt{\sum_{i=1}^{n}(u_{in})^2}} \end{pmatrix} \tag{A-5}$$

Let $o_{ij} = n - o_i/n - o_j$, and the unitised matrix is

$$\bar{O} = (\bar{o}_{ij})_{n\times n} = \begin{pmatrix} \frac{o_{11}}{\sqrt{\sum_{i=1}^{n}(o_{i1})^2}} & \frac{o_{12}}{\sqrt{\sum_{i=1}^{n}(o_{i2})^2}} & \cdots & \frac{o_{1n}}{\sqrt{\sum_{i=1}^{n}(o_{in})^2}} \\ \frac{o_{21}}{\sqrt{\sum_{i=1}^{n}(o_{i1})^2}} & \frac{o_{22}}{\sqrt{\sum_{i=1}^{n}(o_{i2})^2}} & \cdots & \frac{o_{2n}}{\sqrt{\sum_{i=1}^{n}(o_{in})^2}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{o_{n1}}{\sqrt{\sum_{i=1}^{n}(o_{i1})^2}} & \frac{o_{n2}}{\sqrt{\sum_{i=1}^{n}(o_{i2})^2}} & \cdots & \frac{o_{nn}}{\sqrt{\sum_{i=1}^{n}(o_{in})^2}} \end{pmatrix}. \tag{A-6}$$

Unitised multiplicative preference relation matrix is as follows:

$$\bar{A}=\left(\bar{a}_{ij}\right)_{n\times n}=\begin{pmatrix}\frac{a_{11}}{\sqrt{\sum_{i=1}^{n}(a_{i1})^2}} & \frac{a_{12}}{\sqrt{\sum_{i=1}^{n}(a_{i2})^2}} & \cdots & \frac{a_{1n}}{\sqrt{\sum_{i=1}^{n}(a_{in})^2}}\\ \frac{a_{21}}{\sqrt{\sum_{i=1}^{n}(a_{i1})^2}} & \frac{a_{22}}{\sqrt{\sum_{i=1}^{n}(a_{i2})^2}} & \cdots & \frac{a_{2n}}{\sqrt{\sum_{i=1}^{n}(a_{in})^2}}\\ \vdots & \vdots & \ddots & \vdots\\ \frac{a_{n1}}{\sqrt{\sum_{i=1}^{n}(a_{i1})^2}} & \frac{a_{n2}}{\sqrt{\sum_{i=1}^{n}(a_{i2})^2}} & \cdots & \frac{a_{nn}}{\sqrt{\sum_{i=1}^{n}(a_{nn})^2}}\end{pmatrix}. \tag{A-7}$$

Unitised additive preference relation matrix is as follows:

$$\bar{B}=\left(\bar{p}_{ij}\right)_{n\times n}=\begin{pmatrix}\frac{p_{11}}{\sqrt{\sum_{i=1}^{n}(p_{i1})^2}} & \frac{p_{12}}{\sqrt{\sum_{i=1}^{n}(p_{i2})^2}} & \cdots & \frac{p_{1n}}{\sqrt{\sum_{i=1}^{n}(p_{in})^2}}\\ \frac{p_{21}}{\sqrt{\sum_{i=1}^{n}(p_{i1})^2}} & \frac{p_{22}}{\sqrt{\sum_{i=1}^{n}(p_{i2})^2}} & \cdots & \frac{p_{2n}}{\sqrt{\sum_{i=1}^{n}(p_{in})^2}}\\ \vdots & \vdots & \ddots & \vdots\\ \frac{p_{n1}}{\sqrt{\sum_{i=1}^{n}(p_{i1})^2}} & \frac{p_{n2}}{\sqrt{\sum_{i=1}^{n}(p_{i2})^2}} & \cdots & \frac{p_{nn}}{\sqrt{\sum_{i=1}^{n}(p_{nn})^2}}\end{pmatrix}. \tag{A-8}$$

Based on the unitisation of preference relations (A-5 to A-8), $\|\vec{u}_j\|=\|\vec{o}_j\|=\|\vec{a}_j\|=\|\vec{p}_j\|=1$ holds.
Thus, the distances denoted by Definition 2 in Section 4.2 are inner products, and the inner product space is constructed.

Appendix B. Transformation of the dual optimisation model of SVM:

$$L\left(\vec{n},\alpha\right)=\frac{1}{2}\|\vec{n}\|^2-\sum_{i=1}^{n}\alpha_i\left(y_i\left\langle\vec{n},\vec{r}_i\right\rangle-1\right) \tag{B-1}$$

The Lagrange function can be transformed into the following equation using the KKT conditions $\frac{\partial L}{\partial\vec{n}}=0\Rightarrow\vec{n}=\sum_{i=1}^{n}\alpha_i y_i\vec{r}_i$:

$$\begin{aligned}L\left(\vec{n},\alpha\right)&=\tfrac{1}{2}\|\vec{n}\|^2-\sum_{i=1}^{n}\alpha_i\left(y_i\left\langle\vec{n},\vec{r}_i\right\rangle-1\right)\\&=\tfrac{1}{2}\vec{n}^T\vec{n}-\sum_{i=1}^{n}\alpha_i y_i\vec{n}^T\vec{r}_i+\sum_{i=1}^{n}\alpha_i\\&=\tfrac{1}{2}\vec{n}^T\sum_{i=1}^{n}\alpha_i y_i\vec{r}_i-\vec{n}^T\sum_{i=1}^{n}\alpha_i y_i\vec{r}_i+\sum_{i=1}^{n}\alpha_i\\&=-\tfrac{1}{2}\Big(\sum_{i=1}^{n}\alpha_i y_i\vec{r}_i\Big)^T\sum_{i=1}^{n}\alpha_i y_i\vec{r}_i+\sum_{i=1}^{n}\alpha_i\\&=\sum_{i=1}^{n}\alpha_i-\sum_{i,j=1}^{n}\alpha_i\alpha_j y_i y_j\vec{r}_i^T\vec{r}_j.\end{aligned} \tag{B-2}$$

Appendix C. Aggregation of subgroup preferences [2,3]

The perfect consistencies (A-1 to A-4) are difficult to hold in real-life decision-making problems because the perfect consistency condition is not always satisfied due to insufficient experience or knowledge of DMs about alternatives [3,21] and equations (A-1 to A-4) do not always equal to one. The collective opinion is the closest preference to each preference relation based on the similarity measure, that is:

$$\begin{aligned}\text{Max}\quad C=&\sum_{k\in\Omega_U}\sum_{i=1}^{n}\sum_{j=1}^{n}\lambda_k\bar{\omega}_i\bar{u}_{ij}^{(k)}+\sum_{k\in\Omega_O}\sum_{i=1}^{n}\sum_{j=1}^{n}\lambda_k\bar{\omega}_i\bar{o}_{ij}^{(k)}\\&+\sum_{k\in\Omega_A}\sum_{i=1}^{n}\sum_{j=1}^{n}\lambda_k\omega^-{}_i a^-{}_{ij}^{(k)}+\sum_{k\in\Omega_B}\sum_{i=1}^{n}\sum_{j=1}^{n}\lambda_k\omega^-{}_i p^-{}_{ij}^{(k)}\\&S.t.\begin{cases}\sum_{i=1}^{n}\bar{\omega}_i^2=1;\\0\le\bar{\omega}_i\le1.\end{cases}\end{aligned} \tag{C-1}$$

where $\lambda_k=\frac{s_v}{\sum_{i\in M(v)}s_i}$, and $\bar{\omega}=\left(\frac{\omega_1}{\sqrt{\sum_{i=1}^{n}\omega_i^2}},\frac{\omega_2}{\sqrt{\sum_{i=1}^{n}\omega_i^2}},\ldots,\frac{\omega_n}{\sqrt{\sum_{i=1}^{n}\omega_i^2}}\right)^T$. $\bar{u}_{ij}^{(k)},\bar{o}_{ij}^{(k)},\bar{a}_{ij}^{(k)},\bar{p}_{ij}^{(k)}$ is the unitised entries in matrixes (A-5) to (A-8).

Therefore, the collective preference of each subgroup is calculated using the following formula [2]:

$$w_i = \frac{\begin{array}{c}\sum_{k\in\Omega_U}\sum_{j=1}^{n}\lambda_k \overline{u}_{ij}^{(k)} + \sum_{k\in\Omega_O}\sum_{j=1}^{n}\lambda_k \overline{o}_{ij}^{(k)}\\ +\sum_{k\in\Omega_A}\sum_{j=1}^{n}\lambda_k \overline{a}_{ij}^{(k)} + \sum_{k\in\Omega_B}\sum_{j=1}^{n}\lambda_k \overline{p}_{ij}^{(k)}\end{array}}{\sum_{i=1}^{n}\left(\begin{array}{c}\sum_{k\in\Omega_U}\sum_{j=1}^{n}\lambda_k \overline{u}_{ij}^{(k)} + \sum_{k\in\Omega_O}\sum_{j=1}^{n}\lambda_k \overline{o}_{ij}^{(k)}\\ +\sum_{k\in\Omega_A}\sum_{j=1}^{n}\lambda_k \overline{a}_{ij}^{(k)} + \sum_{k\in\Omega_B}\sum_{j=1}^{n}\lambda_k \overline{p}_{ij}^{(k)}\end{array}\right)}, \quad i = 1,2,\ldots,n. \tag{C-2}$$

Appendix D. Initial consensus degree

The consensus degrees in GDM are normally measured by two ways: the similarity of a preference relation at different levels [27] and the deviation of individual ranking to collective ranking of alternatives [15,19]. Cardinal and ordinal consensus degrees are the most frequently used indexes to compare ranking positions of alternatives.

In GDM, a unanimous agreement is difficult to achieve using deviation measures, regardless of 'hard' or 'soft' index [5,18]. However, an acceptable consensus can be achieved if DMs can obtain adequate economic compensations. In this study, a consensus degree (D-3) is defined to show the changes of the total costs for a consensus reaching process with different parameters.

To preset a consensus degree $\xi > 0$, a DM $DM_i^v, i = 1,2,\ldots,m_v$ (m_v is the number of DMs in a subgroup) belongs to a community that is attached to node v in a local social network. The order of alternatives in the derived weights of DM_i^v is $h_i^v = (h_{i1}^v, h_{i2}^v, \ldots, h_{in}^v)$. The collective opinion is $p_c^v = (p_{c1}^v, p_{c2}^v, \ldots, p_{cn}^v)$. The ordinal consensus degree [14] is:

$$OCD(DM_i^v) = \frac{1}{n^2}\sum_{j=1}^{n}\left|h_{ij}^v - p_{cj}^v\right| \tag{D-1}$$

and the consensus degree of subgroup M_v is:

$$OCD(M_v) = \frac{1}{m_v}\sum_{v=1}^{m_v} OCD(DM_v) \tag{D-2}$$

where m_v is the number of DMs in node v.

Definition 5. Consensus Degree: the total consensus degree is the weighted mean of subgroup consensus degree, that is:

$$TOCD(V) = \frac{1}{N} s_v \sum_{v=1}^{m_v} OCD(M_v) \tag{D-3}$$

where N is the number of vertexes in a local social network, and s_v is the weighted centrality degree of each subgroup in Equation (8).

Appendix E. Example

The example comes from Financial Inclusion, which aims to provide financial services at an affordable cost to low-income groups in need. A new alternative method of credit evaluation was developed through a GDM because farmers and herdsmen in extremely poor areas do not have enough financial activity information to evaluate their credit status [4]. Fifty-two participants in this method include local credit unions, village representatives and local government officials. These individuals often communicate with the evaluated objects. Therefore, the evaluation and ranking of the evaluated objects can be carried out through these representatives. The social network relationship often affects these individuals to make their own judgments, including mutual trust relations and individual preferences. In this example, five alternatives will apply an interest-free loan. The five alternatives were provided to the 52 DMs; then, they make their choice and reach consensus through the subgroup partition with minimum cost.

In this example, five alternatives will apply an interest-free loan. The information is listed as follows:

Alternative 1: The potential beneficiary is 55 years old. His family has three laborers. The purpose of the loan is to build a field free-range chicken farm.

Alternative 2: The potential beneficiary is 51 years old single male. The purpose of the loan is to receive a living support when he is out-migration for work.

Alternative 3: The potential beneficiary is 42 years old. His family has two laborers. His family income is agricultural products trading. The purpose of the loan is to obtain a circulating capital for his business.

Alternative 4: The potential beneficiary is 45 years old. His family has four laborers and owns eight cows and 24 sheep. The purpose of the loan is to buy a tractor to improve agricultural production.

Alternative 5: The potential beneficiary is 42 years old and divorced. He needs to take care of his father at home and has a daughter attending high school. The purpose of the loan is to rebuild his house collapsed in heavy rain.

The five alternatives were provided to the 52 DMs, and then they make their choice. The preference relation of the other village representatives (the detailed data have been opened in [4]). However, the trust relation is not considered in their application, and a feedback mechanism is used for consensus reaching. To compare the GDM problems with and without social network, we preset the trust relation amongst part of the representatives who are investigated, loan officers, poverty alleviation administrators, village head and central bank officials. The detailed calculation process is listed as follows:

Step 1: community detection

The 10 representatives with trust relation are detected into three communities using Equation (13). The individual preference and community number are listed as follows:

$$\left(\left|\begin{array}{l} \textit{Serial number} \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \\ 8 \\ 9 \\ 10 \end{array}\right|\right) \overbrace{\begin{array}{ccccc} 0.1427 & 0.3167 & 0.1009 & 0.1231 & 0.3167 \\ 0.0883 & 0.2869 & 0.2105 & 0.2665 & 0.1478 \\ 0.2372 & 0.1210 & 0.2093 & 0.1583 & 0.2742 \\ 0.3015 & 0.1117 & 0.1932 & 0.2179 & 0.1615 \\ 0.3015 & 0.1117 & 0.1932 & 0.2179 & 0.1615 \\ 0.1063 & 0.4159 & 0.0869 & 0.1442 & 0.1607 \\ 0.0922 & 0.1501 & 0.1520 & 0.3819 & 0.1936 \\ 0.1829 & 0.0944 & 0.3731 & 0.0657 & 0.2106 \\ 0.1915 & 0.1850 & 0.3219 & 0.1500 & 0.1683 \\ 0.2565 & 0.2241 & 0.2011 & 0.1544 & 0.2247 \end{array}}^{\textit{Indiviual Preferences(priority vector)}} \left\|\begin{array}{l} \textit{Community number} \\ 2 \\ 1 \\ 1 \\ 1 \\ 1 \\ 2 \\ 1 \\ 3 \\ 3 \\ 2 \end{array}\right\| \tag{E-1}$$

Step 2: subgroup partition

The 52 DMs are divided into three subgroups using SVM in the inner product space (Equation (17)), and the training set and labels are based on Equation (E-1). The remaining DMs will be divided into different groups. The results are listed as follows (see Table E-1):

Table E1
Classification results.

Subgroups	Number of representatives	Size of subgroups	Community preference	Centrality
No. 1	5	31	0.2066 0.1581 0.1939 0.2515 0.1899	0.4395
No. 2	3	15	0.1699 0.3216 0.1307 0.1417 0.2360	0.3001
No. 3	2	8	0.1926 0.1438 0.3576 0.1110 0.1950	0.2604
Total: 3	10	52		

Step 3: weight determination

The adjacent matrix of the three communities is as follows (E-2):

$$A = \begin{array}{c} 1 \\ 2 \\ 3 \end{array}\left(\overbrace{\begin{array}{ccc} 0 & 12 & 10 \\ 12 & 0 & 3 \\ 10 & 3 & 0 \end{array}}^{\textit{Communities}}\right) \tag{E-2}$$

and the similarity of each community is as follows (E-3):

$$D' = (d_{ij}) = \begin{pmatrix} 1 & 0.8935 & 0.9710 \\ 0.8935 & 1 & 0.9623 \\ 0.9710 & 0.9623 & 1 \end{pmatrix} \tag{E-3}$$

This matrix can then be combined with a participant's coefficient vector C by using Equation (1)

$$C = (0.5741 \quad 0.2778 \quad 0.1481)^T \tag{E-4}$$

Then, the $S = [AC^T | D'C^T]$ can be decomposed into:

$$\begin{pmatrix} \theta_1 & \\ & \theta_2 \end{pmatrix} = \begin{pmatrix} 0.9743 & \\ & 0.0257 \end{pmatrix} \tag{E-5}$$

We calculate the weights of the subgroups by using Equation (8), which is the weighted eigenvector centrality of each community:

$$\vec{s} = (s_v) = \begin{array}{c} 1 \\ 2 \\ 3 \end{array} \begin{pmatrix} 0.4395 \\ 0.3001 \\ 0.2604 \end{pmatrix} \tag{E-6}$$

Step 4: optimal ranking of the alternatives with minimum cost

We set the consensus degree $OCD = 0.2113$ because this is the final threshold obtained in [4]. We can obtain the preference modification interval for each community as follows:

$$\xi_v = \begin{pmatrix} 12.0192 \\ 17.6004 \\ 20.2893 \end{pmatrix} \tag{E-8}$$

Thus, we can obtain an optimal collective group opinion of the five alternatives as follows:

$$O_c = (0.2254 \quad 0.1752 \quad 0.2030 \quad 0.2110 \quad 0.1854) \tag{E-9}$$

We can observe the following similarities and differences in the results of the two methods from Table E-2. The first alternative (x_1) and the last alternative (x_2) will not change their ranking during the influence of social networks; however, the intermediate alternatives ($x_3x_5x_4$) will have a ranking change. The weights of the alternatives will be closer through the social network. For example, the difference between the first and the last alternatives is 0.0767, and that between the candidate solutions in the social network is reduced to 0.0502. This notion indicates that opinions between different subgroups are neutralised through the preference similarity classification and trust relationships in social networks, which will help in reducing possible conflicts in the choice of alternatives.

Table E2
Comparisons of the alternatives' ranking.

Order	Feedback mechanism [57]			Our method		
	Alternatives	Weights	Change	Alternatives	Weights	Change
1	x_1	0.2508	–	x_1	0.2254	–
2	x_3	0.2051	↑	x_4	0.2110	↑
3	x_5	0.1900	↑	x_3	0.2030	↓
4	x_4	0.1799	↓	x_5	0.1854	↓
5	x_2	0.1741	–	x_2	0.1752	–

Appendix F. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ins.2021.06.047.

References:

[1] X. Bai, P. Shi, Y. Liu, Society: realizing china's urban dream, Nature 509 (2014) 158–160.
[2] X.R. Chao, G. Kou, Y. Peng, A similarity measure-based optimization model for group decision-making with multiplicative and fuzzy preference relations, International Journal of Computers Communications & Control 12 (2017) 26–40.
[3] X.R. Chao, G. Kou, T. Li, Y. Peng, Jie Ke versus AlphaGo: A ranking approach using decision making method for large-scale data with incomplete information, Eur. J. Oper. Res. 265 (2018) 239–247.
[4] X. Chao, G. Kou, Y. Peng, E.H. Viedma, Large-Scale Group Decision-Making with Non-cooperative Behaviors and Heterogeneous Preferences: An Application in Financial Inclusion, Eur. J. Oper. Res. 288 (2021) 271–293, https://doi.org/10.1016/j.ejor.2020.05.047.
[5] X. Chao, Y. Dong, G. Kou, Y., Peng How to determine the consensus threshold in group decision making: A method based on efficiency benchmark using benefit and cost insight, Ann. Oper. Res. (2021), https://doi.org/10.1007/s10479-020-03927-8.
[6] D. Cheng, Z. Zhou, F. Cheng, Y. Zhou, Y. Xie, Modeling the minimum cost consensus problem in an asymmetric costs context, Eur. J. Oper. Res. 270 (2018) 1122–1137, https://doi.org/10.1016/j.ejor.2018.04.041.
[7] D. Cheng, F. Cheng, Z. Zhou, Y. Wu, Reaching a minimum adjustment consensus in social network group decision-making, Information Fusion 59 (2020) 30–43.
[8] R. Cialdini, The science of persuasion, Sci. Am. 284 (2001) 76–81.
[9] I.D. Couzin, C.C. Ioannou, G. Demirel, T. Gross, C.J. Torney, A. Hartnett, et al, Uninformed individuals promote democratic consensus in animal groups, Science 334 (2011) 1578.
[10] Z. Ding, X. Chen, Y. Dong, F. Herrera, Consensus reaching in social network DeGroot Model: The roles of the Self-confidence and node degree, Inf. Sci. 486 (2019) 62–72.
[11] L. Donetti, M.A. Munoz, Detecting network communities: a new systematic and efficient algorithm, J. Stat. Mech: Theory Exp. 10 (2004) 10–12.
[12] Y. Dong, Q. Zha, H. Zhang, G. Kou, H. Fujita, F. Chiclana, E. Herrera-Viedma, Consensus reaching in social network group decision-making: Research paradigms and challenges, Knowl.-Based Syst. 162 (2018) 3–13.
[13] Y. Dong, H. Zhang, E. Herrera-Viedma, Integrating experts' weights generated dynamically into the consensus reaching process and its applications in managing non-cooperative behaviors, Decis. Support Syst. 84 (2016) 1–15.

[14] Y. Dong, H. Zhang, Multiperson decision-making with different preference representation structures: a direct consensus framework and its properties, Knowl.-Based Syst. 58 (2014) 45–57.
[15] S. Fortunato, D. Hric, Community detection in networks: A user guide, Phys. Rep. 659 (2010) 1–44.
[16] Z. Gong, H. Zhang, J. Forrest, L. Li, X. Xu, Two consensus models based on the minimum cost and maximum return regarding either all individuals or one individual, Eur. J. Oper. Res. 240 (2015) 183–192.
[17] M. Gupta, Consensus building process in group decision making—An adaptive procedure based on group dynamics, IEEE Trans. Fuzzy Syst. 26 (4) (2017) 1923–1933.
[18] E. Herrera-Viedma, F.J. Cabrerizo, J. Kacprzyk, W. Pedrycz, A review of soft consensus models in a fuzzy environment, Information Fusion 17 (2014) 4–13.
[19] E. Herrera-Viedma, F. Herrera, F. Chiclana, A consensus model for multiperson decision-making with different preference structures, IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans 32 (2002) 394–402.
[20] F. Jing, X. Chao, Fairness Concern: An Equilibrium Mechanism for Consensus-Reaching Game in Group Decision-Making, Information Fusion 72 (2021) 147–160.
[21] G. Kou, C. Lin, A cosine maximization method for the priority vector derivation in AHP, Eur. J. Oper. Res. 235 (2014) 225–232.
[22] G. Kou, Y. Peng, X. Chao, E. Herrera-Viedma, F.E. Alsaadi, A geometrical method for consensus building in GDM with incomplete heterogeneous preference information, Appl. Soft Comput. (2021), https://doi.org/10.1016/j.asoc.2021.107224.
[23] Á. Labella, Y. Liu, R.M. Rodríguez, L. Martínez, Analyzing the performance of classical consensus models in large-scale group decision-making: a comparative study, Appl. Soft Comput. 67 (2018) 677–690, https://doi.org/10.1016/j.asoc.2017.05.045.
[24] B. Liu, Q. Zhou, R.X. Ding, I. Palomares, F. Herrera, Large-scale group decision-making model based on social network analysis: Trust relationship-based conflict detection and elimination, Eur. J. Oper. Res. 275 (2019) 737–754.
[25] Y. Liu, T. Zhou, J.Y.L. Forrest, A Multivariate Minimum Cost Consensus Model for Negotiations of Holdout Demolition, Group Decis. Negot. (2020), https://doi.org/10.1007/s10726-020-09683-1.
[26] L.V.D. Maaten, G. Hinton, Visualizing Data using t-SNE, Journal of Machine Learning Research 9 (2008) 2579–2605.
[27] I. Palomares, L. Martinez, F. Herrera, A consensus model to detect and manage noncooperative behaviors in large-scale group decision-making, IEEE Trans. Fuzzy Syst. 22 (2014) 516–530.
[28] I.J. Pérez, F.J. Cabrerizo, E. Herrera-Viedma, A mobile decision support system for dynamic group decision-making problems, IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans 40 (2010) 1244–1256.
[29] J.A. Recio-Garcia, L. Quijano, B. Diaz-Agudo, Including social factors in an argumentative model for group decision support system, Decis. Support Syst. 56 (2013) 48–55.
[30] V. Roquet, L. Bornholdt, K. Sirker, and J. Lukic. Urban Land Acquisition and Involuntary Resettlement Linking Innovation and Local Benefits. The World Bank. This work is available under the Creative Commons Attribution 3.0 IGO license (CC BY 3.0 IGO) http://creativecommons.org/licenses/by/3.0/igo., 2017.
[31] T.L. Saaty, The Analytic Hierarchy Process, McGraw-Hill, New York, USA, 1980.
[32] W. Stanley, F. Katherine, Social Network Analysis in the Social and Behavioral Sciences. Social Network Analysis: Methods and Applications. Cambridge University Press. pp. 1–27. ISBN 9780521387071.1994.
[33] R. Ureña, F. Chiclana, J.A. Morente-Molinera, E. Herrera-Viedma, Managing incomplete preference relations in decision-making: a review and future trends, Inf. Sci. 302 (2015) 14–32.
[34] R. Urena, F. Chiclana, G. Melancon, E. Herrera-Viedma, A social network-based approach for consensus achievement in multiperson decision-making, Information Fusion 47 (2019) 72–87.
[35] R. Urena, G. Kou, Y. Dong, F. Chiclana, E. Herrera-Viedma, A review on trust propagation and opinion dynamics in social networks and group decision-making frameworks, Inf. Sci. 478 (2019) 461–475.
[36] V.N. Vapnik, An overview of statistical learning theory, IEEE Trans. Neural Networks 10 (1999) 988–999.
[37] J. Wu, F. Chiclana, H. Fujita, E. Herrera-Viedma, A visual interaction consensus model for social network group decision-making with trust propagation, Knowl.-Based Syst. 122 (2017) 39–50.
[38] J. Wu, F. Chiclana, A social network analysis trust–consensus based approach to group decision-making problems with interval-valued fuzzy reciprocal preference relations, Knowl.-Based Syst. 59 (2014) 97–107.
[39] J. Wu, Y. Yang, P. Zhang, L. Ma, Households' noncompliance with resettlement compensation in urban China: Toward an integrated approach, International Public Management Journal 21 (2018) 272–296.
[40] T. Wu, K. Zhang, X. Liu, C. Cao, A two-stage social trust network partition model for large-scale group decision-making problems, Knowl.-Based Syst. 163 (2019) 632–643.
[41] T. Wu, X. Liu, F. Liu, An interval type-2 fuzzy topsis model for large scale group decision making problems with social network information, Inf. Sci. 478 (2019) 461–475.
[42] Z. Wu, J. Xu, A consensus model for large-scale group decision-making with hesitant fuzzy information and changeable clusters, Information Fusion 41 (2018) 217–231.
[43] X.H. Xu, Z.J. Du, X.H. Chen, Consensus model for multi-criteria large-group emergency decision-making considering non-cooperative behaviors and minority opinions, Decis. Support Syst. 79 (2015) 150–160.
[44] M. Yang, G. Adomavicius, G. Burtch, Y.C. Ren, Mind the gap: accounting for measurement error and misclassification in variables generated via data mining, Information systems research 29 (2018) 4–24.
[45] H. Zhang, G. Kou, Y. Peng, Soft consensus cost models for group decision making and economic interpretations, Eur. J. Oper. Res. 277 (2019) 964–980.
[46] H. Zhang, Y. Dong, F. Chiclana, S. Yu, Consensus efficiency in group decision making: A comprehensive comparative study and its optimal design, Eur. J. Oper. Res. 275 (2) (2019) 580–598.