sustainability

MDPI

Article

A Spatial-Territorial Reorganization Model of Rural Settlements Based on Graph Theory and Genetic Optimization

Yan Mao [1], Yanfang Liu [1,2,*], Haofeng Wang [1], Wei Tang [1] and Xuesong Kong [1]

1 School of Resource and Environment Science, Wuhan University, 129 Luoyu Road, Wuhan 430079, China; maoyan@whu.edu.cn (Y.M.); hhfwang2009@whu.edu.cn (H.W.); tangwei@geoway.com.cn (W.T.); xuesongk@whu.edu.cn (X.K.)

2 Key Laboratory of Geographic Information System, Ministry of Education, Wuhan University, Wuhan 430079, China

* Correspondence: yfliu6106@163.com; Tel.: +86-27-6877-8893

Received: 26 June 2017; Accepted: 28 July 2017; Published: 3 August 2017

Abstract: Rural China has experienced rapid urbanization and industrialization, accompanied with rural–urban migration since 1978. This tremendous transition has caused a series of negative consequences, necessitating a spatial-territorial reorganization of rural settlements. Previous studies on the restructuring of rural settlements are insufficient for inter-settlement connection consideration and practical and dynamic decision-making techniques. To overcome these concerns, a dynamic spatial-territorial reorganization model (SRM) of rural settlement is proposed herein based on graph theory and genetic algorithm (GA). The model involves two parts. In Part 1, consolidated settlements are identified according to the socio-economic network performance under four types of attack. In Part 2, GA model is repeatedly executed to scientifically resettle consolidated settlements into nearby townships or central settlements with objectives of suitability, compactness, and local connectivity under the control of the constraints. This paper presents an application of SRM to Chengui Town, Hubei Province. Empirical results suggest that: (1) removing settlements in order of node degree is the least efficient way to destroy the entire functional system; and (2) the proposed model can yield satisfactory solutions in terms of spatial reorganization of settlements. The SRM may also serve as a valuable reference for planners in devising plans and making decisions.

Keywords: spatial-territorial reorganization; rural settlements; graph theory; genetic algorithm; local connectivity; robustness

1. Introduction

Since 1978, rural China has experienced rapid urbanization and industrialization, accompanied with rural–urban migration [1–3]. Owing to this tremendous transition, in China a unique rural settlement morphology has emerged; that is, hollowed villages [4]. The "hollowed villages" is a phenomenon of depopulation leading to abandonment of buildings and land in rural communities, due to the dual-track structure of rural–urban development (i.e., urban land is state owned, whereas farmland is collectively owned) and the restriction of hukou (i.e., household registration system) [4,5]. Such phenomenon has caused several negative consequences, such as the weakening function of critical rural organizations, the fragility of structures and networks, the chaotic flow of rural development elements, and the lack of economies of scale and output efficiency [5,6]. Hence, the current layout of China's rural settlements, which can be described as "scattered, massy, small, and hollowed," demands emergency measures. For example, the consolidation of the hollowed villages primarily aims to promote the spatial-territorial reorganization, which is accompanied with administrative reorganization in rural restructuring [1,4]. This type of rural restructuring encourages concentrating the rural population in communities or central settlements and merging settlements [1,4]. Nevertheless, the majority of regional studies are concerned with urban studies or theoretical research on rural settlements (e.g., policy analysis, strategic decision and planning) [1,3,5–7]. Existing studies on rural settlement restructuring remain insufficient in terms of two aspects: consideration of inter-settlement connections and absence of dynamic and practical decision-making techniques. To address these research gaps, we propose a dynamic spatial-territorial reorganization model (SRM) of rural settlements based on graph theory and genetic optimization.

The traditional village system consists of relatively independent villages or rural settlements [7]. Recently, inter-settlement interactive scope and content are expanding with population as well as social, economic, and traffic flows [7–9]. The spatial-territorial reorganization or restructuring of rural settlements (involving settlement removal and incorporation) should consider inter-settlement connections [10]. Moreover, this consideration should entail two aspects. First, consolidated settlements should exert the least effect on the entire village system to maximally maintain the functionality of system and the stability of the villagers' life when they are removed. Second, consolidated settlements should be relocated to adjacent high-related central settlements to reduce the separation between villagers after consolidation. In response to these requirements, the robustness of network and local connectivity are introduced in this paper. Similar to other complex systems, inter-settlement connection system can be modeled as a network, in which settlements are nodes and interactions and activities among settlements are edges [9,11,12]. The robustness of inter-settlement network refers to its ability to maintain the functionality under attacks or failures [11]. Accordingly, settlement importance can be evaluated based on how much the removal of the node "disrupts" the graph structure [11,13,14]. Specifically, if removing a settlement exerts no noticeable effect on the network structure, the settlement is a good option to serve as a settlement to be absorbed or relocated. Local connectivity represents the frequency of all types of connections and activities (working, visiting, shopping, and entertainment) from consolidated settlements to central settlements. Local connectivity acts as an important objective in the SRM to search for high-related relocated settlements in the following sections. This paper provides a system perspective to realize the reorganization of rural settlements rather than a simple individual analysis [6,10].

Existing studies on rural settlement restructuring mainly concern theoretical research, policy analysis, and macro and static planning [1,3,5–7]. Practical and dynamic decision-making techniques are in demand to scientifically realize reorganization and optimization. Several decision-making techniques have been proposed for land-use planning [15–17]. In particular, most of them make use of linear programming when a single clear objective or even multi-objective problems can be identified [17]. Although the linear programming models can quickly lead to optimal solutions [16,17], they cannot cope with large combinatorial optimization problems within reasonable time [15,18,19], incommensurable and/or conflicting objectives [19], and spatial optimization [16,20]. To overcome these concerns, various heuristic algorithms have been developed, such as simulated annealing algorithm [15,21], particle swarm algorithm [22], and genetic algorithm (GA). GA, as introduced by Holland [23] and described in detail by Goldberg [24], optimizes by mimicking the genetic procedures of natural selection and reproduction observed in populations for adaptation and survival. As one of the most robust heuristics [25], GA has been applied to provide optimization solutions for different spatial optimization problems, such as land-use planning [16,18,25,26], optimal location search [15], forestry management [19], urban planning [27], and water allocation planning [28], and confirmed effective. To our knowledge, such an approach has rarely been used in the research on the reorganization and optimization of rural settlements. This paper therefore presents a GA to dynamically realize the reorganization of rural settlements.

Combining the above two points, this paper provides a spatial-territorial modeling technology of rural settlements based on graph theory and GA. The SRM is expected to solve the existing land use problems of rural settlements (e.g., scattered, massy, small, and hollowed). This approach may serve as a valuable reference for planners in devising plans and making decisions. Section 2 describes the spatial-territorial reorganization in detail. Section 3 provides the details of the SRM. Section 4 introduces the study area and relevant data. Section 5 describes and analyzes the results, and the final part gives conclusions.

2. Spatial-Territorial Reorganization of Rural Settlements

Spatial-territorial reorganization is a type of rural restructuring in which villagers are encouraged to relocate in communities or central settlements and incorporate settlements [1]. Spatial-territorial reorganization aims to rejuvenate dispersed, abandoned, and idle rural settlements to improve the effectiveness and efficiency of local governance in rural communities and restructure suitable living space [4]. Long et al. [4] reported three modes of spatial-territorial reorganization depending on geographical context. These modes are: (1) settlement to city, to incorporate urban and peri-urban settlements to cities; (2) settlement to township, to resettle villagers to nearby small towns; and (3) settlement to settlement, to relocate residents from scattered settlements to central settlements (Figure 1). Accordingly, the problem of spatial-territorial reorganization involves two main parts: identification of settlements to be relocated (i.e., which settlements need consolidation) and consolidated settlement relocation (i.e., where to reallocate these settlements).

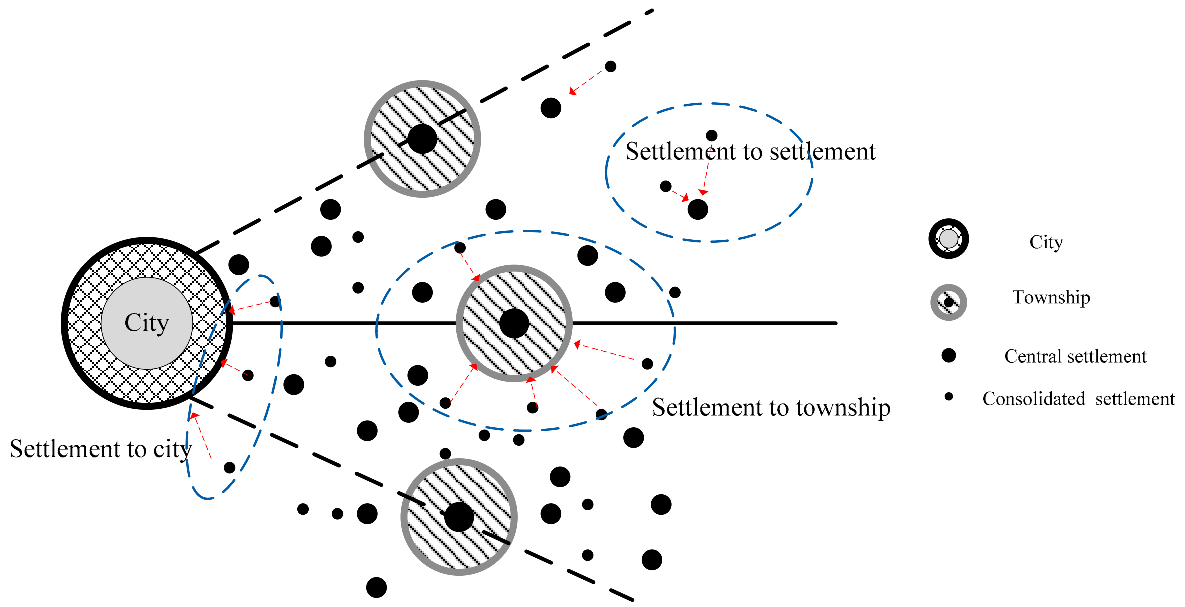

Figure 1. Conceptualization of the three modes of spatial-territorial reorganization.

3. Specifications of SRM of Rural Settlements

SRM is a modeling technology for generating compact and contiguous settlement pattern based on graph theory and GA. The model involves two parts (Figure 2). In Part 1, consolidated settlements are identified according to the network performance under attack. In Part 2, GA model, equipped with objectives of suitability (S), compactness (C), and local connectivity (L) under the control of the constraints, is repeatedly executed to scientifically resettle consolidated settlements into cities, nearby townships, or central settlements.

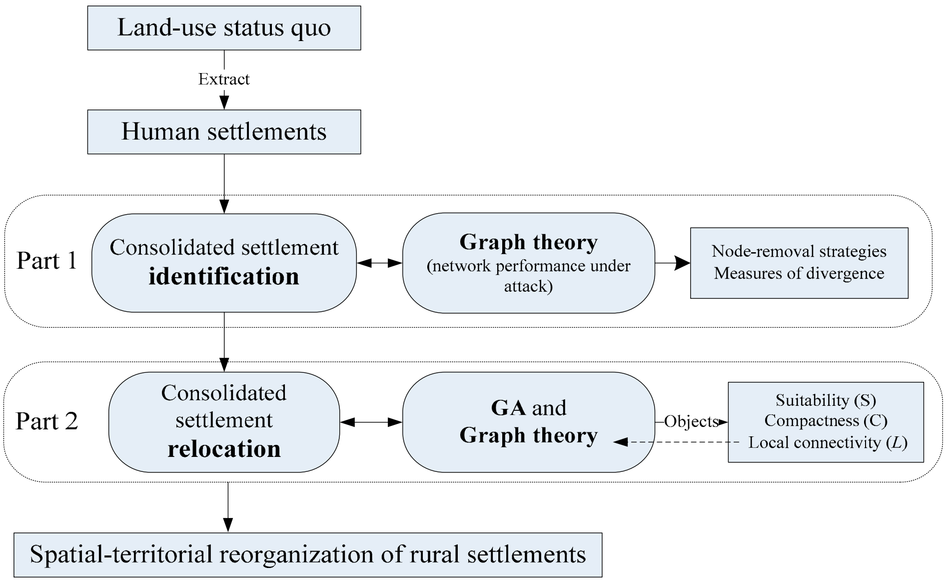

Figure 2. Model framework.

3.1. Consolidated Settlement Identification

To identify consolidated settlements with the least effect on the village system, we generate a series of experiments to investigate the network performance under numerous removals of settlements. First, we design a certain order of settlements. Next, we successively remove settlements following the corresponding strategy order until the network only has one settlement. Finally, the reactions and divergences of the network after each attack are recorded. If the divergence increases slowly, then the corresponding strategy is treated as a reliable node-importance measure.

3.1.1. Node-Removal Strategies

We consider two types of hypothetical node-removal strategies: random and targeted attacks. For targeted strategies, the two most widely used node-importance characters, namely, degree and betweenness, are chosen in this paper [11]. We also use the single node-importance order evaluated by single attack as our fourth strategy to figure out the relationship between single and successive attacks.

Random. No strategy; we randomly choose attacked settlements.

Minimum-degree first. We remove nodes in increasing in-degree order. Degree is defined as the number of connections of a given node [12,14]. A high degree demonstrates that the node connects to more nodes [11]. Boldi, Rosa and Vigna [14] reported that this strategy is a baseline as the degree is the first shot at centrality in a network.

Minimum-betweenness first. We remove node in increasing betweenness order. The betweenness index is computed by identifying the shortest paths linking pairs of nodes and counting the number of times these paths cross each Freeman node [11,12]. Betweenness is a type of global factor [11,12].

Minimum single node-removal importance first. This strategy includes two steps. First, we randomly remove a settlement at each step and obtain the node-removal importance rank according to the record after single attack. Then, we successively remove settlements in increasing single node-removal importance order until the network only has one settlement.

3.1.2. Measures of Divergence

Previous literature often used the diameter or analogous measures to establish whether the network structure has significantly changed after deleting some nodes [13,14,29]. In this paper, we choose six widely used measures to evaluate the network structure: the number of directly reachable pairs *(RP)*, the number of indirectly reachable pairs *(IRP)*, the average length of the shortest paths *(l)*, the cluster coefficient *(C)*, the number of components *(NC)*, and the relative size of the largest connected component *(S).* Boldi, Rosa and Vigna [14] believed that the *RP* and *IRP* are the most immediate global features that are computationally approachable. To evaluate the efficiency behavior of network under attack from both global and local viewpoints, l and C are used [11]. *NC* and S are the most efficient means to measure network fragmentation under failures [11].

The number of directly reachable pairs. It is the number of settlement pairs $<x, y>$ which have a direct path from settlement x to y [14].

The number of indirectly reachable pairs. It is the number of pairs $<x, y>$ which have a path from settlement x to y through other settlements.

The average length of the shortest paths. It measures the separation between settlements in a network.

$$l = \frac{1}{n(n-1)} \sum_{v \in v} \sum_{w \neq v \in v} d(v, w) \tag{1}$$

where n is the number of settlements, and $d(v, w)$ denotes the length of the shortest path between settlements v and w.

The cluster coefficient. It is defined as the probability that two settlements are directly connected to a third settlement and each other.

$$c = \frac{1}{n} \sum_{v \in v} c_v \tag{2}$$

where n is the number of settlements, and C_V is the clustering coefficient of settlement v.

$$c_v = \frac{|\{e_{uw} : u, w \in N_i, e_{uw} \in E\}|}{k_v(k_v - 1)/2} \tag{3}$$

where N_i is the set that contains all neighbors of settlement v, and k_v is the size of set N_i. Thus, $k_v(k_v - 1)/2$ is the maximal number of edges between k_v settlements, and $|e_{uw}|$ is the number of edges existing in the network between all k_v neighbor of settlement v. E is all edges existing in the network. A higher C indicates a more clustered neighborhood [12].

The number of components. Components are sub-graphs that are connected within but disconnected among sub-graphs [30]. This measure is a helpful tool to reveal the intermediate scales of network organization [12].

The relative size of the largest connected component. S is defined as:

$$s = \frac{n_s}{n} \tag{4}$$

where n_s is the number of settlements in the largest connected component, and n is the number of settlements in the initial network.

Once the metrics that can reflect the network structure are determined, the divergence of these metrics can be evaluated by relative change r [14,29].

$$r(x) = 1 - \frac{x_{new}}{x_{original}} \tag{5}$$

where $x_{original}$ is the metric value in the original network, and x_{new} is the metric value after node removal.

3.2. Consolidated Settlement Relocation

3.2.1. Chromosome Representation

The most common representation in land-use optimization is the vector and grid chromosome [27]. To effectively represent the land use and conveniently manipulate the land units, we select the grid chromosome. The problem of spatial-territorial reorganization of rural settlements can be defined as a search for the most suitable central settlements for consolidated settlements within a spatial dimension ($M \times N$ grids). Accordingly, we introduce the concept of "source" and "sink" to the genotype representation (Figure 3). The "source" represents consolidated settlements, and the "sink" denotes central settlements. All grid cells of source have a value of 1; all sink cells have a value of 2; and others are 0. All settlement patches are encoded. A patch is a set of cells that are allocated to the same land-use type [16]. Grid cells in a settlement patch share the same number. The chromosome has m genes, and m is the number of sources. Each gene represents a relocation strategy for a source.

In rural China, farmers are limited to selling their poor-returning farmland due to the restrictions on the sale of property and ambiguous property rights [4,31]. To protect the farmland, we apply migration radius d during reorganization. The central grid cell of each source can be calculated by Equations (6) and (7). The final migration zone of each source can be identified by four values: $I - d$, $I + d$, $J - d$, and $J + d$ (Figure 3a). The migration radius d will increase in multiples if no central settlement exists in the current migration region. The central settlements within the migration region form an alternative solution data set S of each source (e.g., the S of source 3 has two alternatives: 1 and 4). Each source randomly chooses a sink from S as initialization value. Once all solutions are founded, sources will be incorporated to corresponding sinks with a scale coefficient α ($0 < \alpha < 1$, the new area A_2 of source equals the original area A_1 multiplied by α) (Figure 3b).

$$I = \left\lfloor \frac{\sum_{n=1}^{N} i_n}{N} \right\rfloor \tag{6}$$

$$J = \left\lfloor \frac{\sum_{n=1}^{N} j_n}{N} \right\rfloor \tag{7}$$

where I and J are the row and column number of central grid cells, respectively; N is the number of grid cells in a source; i_n is the row number of grid cell n; and j_n is the column number of grid cell n.

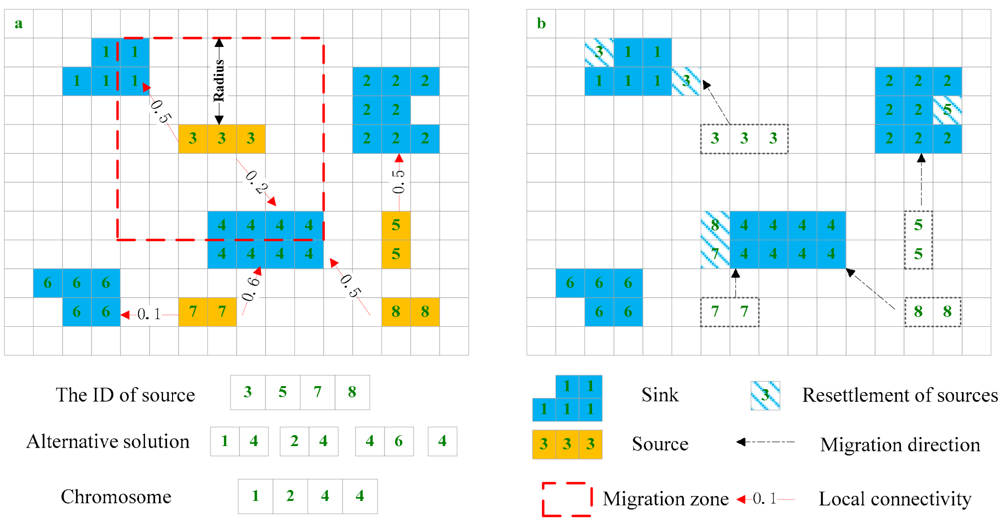

Figure 3. Chromosome structure (**a**); and the resettlement of sources (**b**).

3.2.2. Objectives

Suitability S, compactness C, and local connectivity L are the objectives considered in the SRM. Improving the settlement suitability is conducive to the rational use of land resource [16,32]. Compactness is a basic desirable feature in land use management [28]. The landscape shape index is chosen as the measure for land-use compactness [16]. The objective of local connectivity is an effective way to comprehensively consider social and economic connections between source settlements and sinks in the relocation. Local connectivity represents the frequency of all types of activities (working, visiting, shopping, and entertainment) from consolidated settlements to central settlements. This objective is based on the idea that high-related relocation (i.e., relocate sources to adjacent high-related central settlements) will be helpful for the integration of villagers between sources and sinks after resettlement (Figure 3b). To satisfy these objectives, we integrate the three objectives by a weighted sum. The performance of each solution or chromosome is assessed by the fitness function, as shown in Equation (11). The values of three objectives are normalized by Equation (12).

$$S = \frac{\sum_{i=1}^{M}\sum_{j=1}^{N} s_{ij} * u_{ij}}{\sum_{i=1}^{M}\sum_{j=1}^{N} u_{ij}} \tag{8}$$

$$C = \sum_{h=1}^{H} \frac{P_h}{4\sqrt{A_h}} \tag{9}$$

$$L = \sum_{t=1}^{T} l_t \tag{10}$$

$$Maximize : F = w_1 * f_{norm}^{s}(S) + w_2 * f_{norm}^{c}(C) + w_3 * f_{norm}^{L}(L) \tag{11}$$

$$f_{norm}^{s}(S) = \frac{S - S_{\min}}{S_{\max} - S_{\min}},\ f_{norm}^{c}(C) = \frac{C_{\min}}{C},\ \text{and}\ f_{norm}^{L}(L) = \frac{L - L_{\min}}{L_{\max} - L_{\min}} \tag{12}$$

where s_{ij} is the suitability of rural settlements in the cell indexed by i and j. u_{ij} is a binary-state variable that is 1 if the cell indexed by i and j is located in rural settlement; otherwise, the value is 0. In Equation (9), H is the number of rural settlement patches in the chromosome, P_h is the perimeter of a patch, and A_h is the area of a patch. A small value of C indicates the chromosomes with compact spatial pattern of settlements. In Equation (10), L is defined as the sum of the weights of all edges from source nodes to sink nodes in a network after processing the SRM. T is the number of source patches, l_t is the local connectivity value from source t to corresponding sink after consolidation or relocation. $f_{norm}^{s}(S)$, $f_{norm}^{c}(C)$ and $f_{norm}^{L}(L)$ are normalized variables regarding objective S, objective C, and objective L, and w_1, w_2 and w_3 are the weights of these variables. These weights represent the preference of each variables [15], and they should satisfy the following constraints: (1) $w_1 + w_2 + w_3 = 1$; and (2) $0 \leq w_1 \leq 1$, $0 \leq w_2 \leq 1$, and $0 \leq w_3 \leq 1$. $S_{\max}$ and $L_{\max}$ are the ideal values for objective S and objective L, and $S_{\min}$, $C_{\min}$ and $L_{\min}$ are the worst values for each objective. This paper only uses the worst value $C_{\min}$ to normalize the objective C due to that it is challenging to estimate the ideal value of objective C (see the study of Liu et al. [16] for details).

3.2.3. Constraints

Normally the constraints can be divided into two types: (1) area constraints; and (2) land-use constraints [16]. The area constraints aim to macroscopically maintain a rational land use structure in optimization, such as the maximum and minimum number of cells for different land uses [16]. The land-use constraints restrict land-use conversion within specific grid cells. Here, we focus primarily on the land-use constraints. Especially, in SRM, land in basic farmland zones, reserved green open spaces, or other high-cost conversion areas (e.g., water, industrial and mining land) should restrict the conversion to settlements. Moreover, each cell can only have one land use type [28]. Within the model, the land-use constraints are considered by a restricted layer. In the restricted layer, the restricted land is encoded as 1 and the other land is 0. Only the land with code 0 can be converted to settlements in the process of resettlement of sources.

3.2.4. Genetic Operators

Roulette-wheel technique is applied as the selection operator in our SRM [33]. A single-point crossover with random point-selection is implemented [27]. Specifically, we randomly cut two parent chromosomes right after the crossover gene and exchange the genes following the crossover gene with the crossover probability P_c [32]. The "mutation" operator changes each one of the genes with a probability P_m, i.e., the gene $g(i)$, to be mutated, loses its current value, and receives a random value from data set S depending on the part of the chromosome to which it belongs.

4. Study Area and Data Sources

4.1. Study Area

Chengui Town (114°43′ to 114°49′ E, 30°03′ to 35°30′ N) is located in the Hubei Province, middle of China (Figure 4). In 2014, the town has a total area of 160.40 km^2 and a population of 66,185 people. Chengui has pioneered the province in economic development. Its administrative hierarchy is town (township) and village, and the corresponding settlements are market town and rural settlements. Land-use data in 2014 in vector format were obtained from local government through Huangshi's land consolidation planning project. One township (Chengui township) and 335 rural settlements patches were identified in the vector map (Figure 4). The total area of rural settlement reaches 1073.55 hm^2, and it is inhabited by a rural population of 60,727. Per capita rural settlement land is approximately 176.78 m^2, which is 26.78 m^2 above the national standard (defined by the National Village and Township Planning Code). The land-use data were converted to a 1225 × 1797 grid with a resolution of 10 × 10 m.

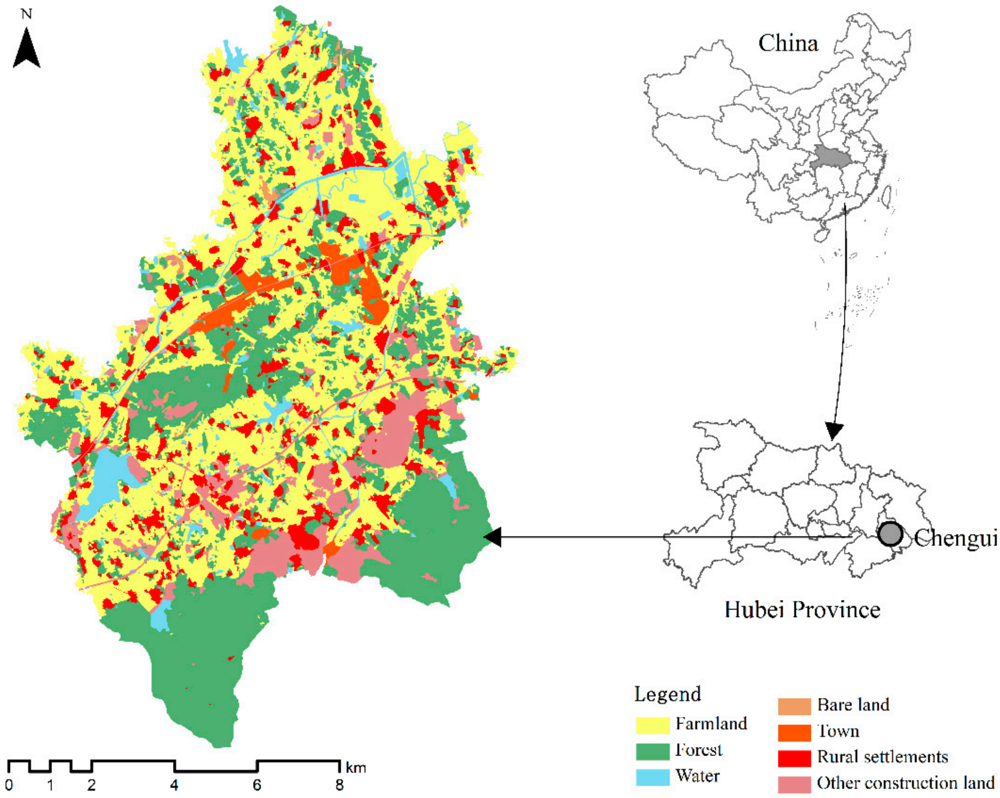

Figure 4. Location and land use of the study area, Chengui.

4.2. Data Resources

Information about socio-economic ties among 336 settlements came from face-to-face interviews within Chengui in autumn 2014. In each settlement, selected questions were asked about the relationships between settlements. For instance, "How many people and households live in this settlement?", "Where do you go to work, attend school, recreate, visit or something else?", and "How many times do you go to that settlement or place in a year?" Based on these questions, we could identify whether villagers tended to move from the responding settlements to other settlements and, if so, how often. During the survey, ties with eight settlements outside of Chengui were also noted. A total of 2085 respondents (3.15% of all population) were survey using a random sampling method. The total number of valid questionnaires was 2056, with 55 respondents in Chengui township. Overall, 2001 respondents were selected in the 335 rural settlements. The average number of respondents in a rural settlement was six, with a sample density (the proportion of respondents to total rural population) of 3.30%. During the survey, we tried to categorize relationships according to different activity types (working, visiting, shopping, and entertainment). Unfortunately, only 454 questionnaires covered this. Therefore, socio-economic ties represented all types of activities among settlements. A given pair of settlements had multiple frequency values due to more than 1 questionnaire in a settlement. To address this, we used mean values to represent the flows among settlements. Then, we standardized the frequency of ties by dividing it by 365. The final values were on a scale of 0–1 (daily = 1, no directed edge connecting = 0). Finally, a weighted adjacency matrix A was constructed. The weights w_i described the frequency of socio-economic activities between settlements *i* and j. We then imported the matrix as a data layer into ARCGIS 10.2 (Figure 5). The lines with arrows indicate the presence of frequent socio-economic interactions between settlements.

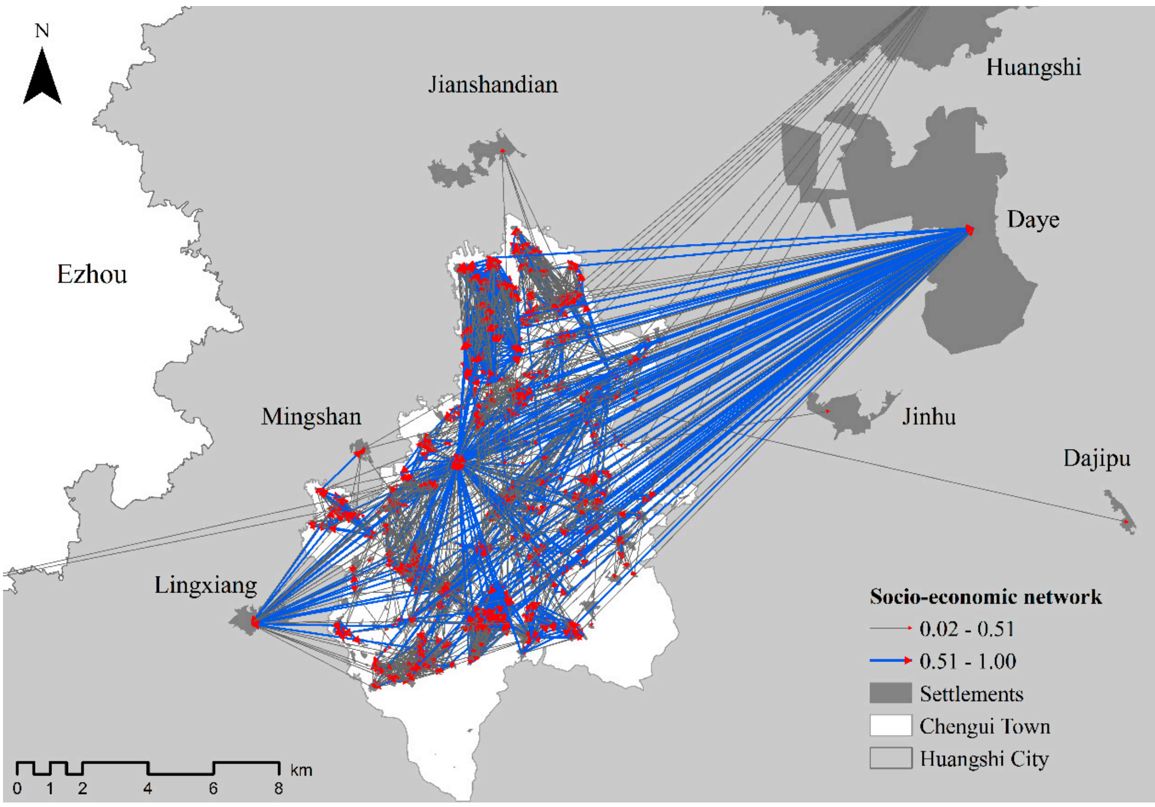

Figure 5. Socio-economic network.

5. Results and Discussion

5.1. Consolidated Settlement Identification

We initially measured settlement importance based on the performances of network under different failures. Through this approach, we can find satisfied consolidated settlements with the least effect on the entire village system. This part involves two steps: evaluating the network structure changes under single removal and assessing network reactions under four successive attacks. The result of single attack is shown in Table 1. The single strategy provides little guidance on the identification due to the minor changes in the network structure. Note that *IRP* suffers relatively dramatic changes. We suppose that the fluctuation is largely due to the exaggeration and overlap of the damage according to the definition of *IRP.*

Figure 6 shows the structural changes of socio-economic network in function of the number of removed settlements under four successive attacks. For ease of interpretation, we also calculated relative change r of *RP*, *IRP*, *L*, and *C* when half of settlements were removed (we did not report the relative change r of *NC* and *S* because these metrics concentrate on the final network structure rather than the relative change) (Table 2). According to Figure 6 and Table 2, our major findings are as follows. First, the network is relatively stable against degree-based and betweenness-based attacks but fragile to random and single node-removal importance attacks. In detail, the random strategy and single node-removal importance strategy disconnect nearly half of direct ($r(RP_{random}) = 44\%, r(RP_{\mathrm{Single}}) = 47\%$) or 70% of indirect pairs ($r(IRP_{random}) = 76\%, r(IRP_{\sin gle}) = 71\%$) of network by removing 50% of settlements, whereas two centrality-based strategies disconnect only roughly 30% of direct ($r(RP_{\mathrm{degree}}) = 26\%, r(RP_{betweenness}) = 51\%$) or 50% of indirect pairs ($r(IRP_{\mathrm{degree}}) = 41\%, r(IRP_{betweenness}) = 49\%$). As for *L* and *C*, the results of all strategies remain stable even when 30% of settlements are removed. When f gets larger (i.e., half of settlements are removed), two centrality-based strategies show higher stability especially for betweenness-based strategy ($r(L_{random}) = 15.87\%, r\left(L_{\sin gle}\right) = 16.38\%, r\left(L_{\mathrm{degree}}\right) = 11.58\%, r(L_{betweenness}) = 8.06\%$, $r(C_{random}) = -21.95\%, r(C_{\sin gle}) = -3.03\%, r(C_{\mathrm{degree}}) = -13.99\%, r(C_{betweenness}) = 15.17\%$). Regarding *NC*, centrality-based strategies consistently attack isolated settlements or small components, whereas random and single node-removal importance strategies continually damage large components. This observation is based on the finding that the network under centrality-based failures keeps two large components even when 28 settlements are removed, whereas the *NC* of random and single node-removal importance strategies changes all the time. As for *S*, centrality-based strategies can also achieve enhanced results ($S_{random} = 42.44\%, S_{\sin gle} = 41.28\%, S_{\mathrm{degree}} = 49.71\%, S_{betweenness} = 50.19\%$ by removing half of settlements). Accordingly, centrality-based strategies can efficiently identify structurally unimportant or important settlements in a socio-eonomic network. The low efficiency of single node-removal importance strategy suggests that the removal of multiple nodes is a complex problem rather than a simple effect superposition of single node-removal. The second important observation is that two centrality-based strategies are similar and diveregent with each other. On one hand, both of them show high efficient to maintain the functionality of the network under attacks. The rank provided by degree is highly correlated to betweenness rank in our empirical data (r = 0.77). On the other hand, degree-based strategy is good at maintaining the accessibility and connectivity of network (*RP* and *IRP*), whereas betweenness-based strategy shows high potential in the efficiency behavior of network (*L* and *C*). The abnormal situation in *L* and *C* (i.e., significantly decrease in *L* and increase in *C*) using random, single node-removal, or degree-based strategies is largely caused by the separation and fragmentation of network (i.e., the network splits into several clusters). Considering the good performance on the basic raw datum (i.e., *RP* and *IRP*) and the stability of *L* and *C* in the first half of f, the degree centrality is chosen as the reliable node-removal importance measure in our settlement study.

Table 1. Performance of socio-economic network under a single attack.

Metrics	Mean	Standard Deviation	Minimum	Median	Maximum
RP	3136.01	16.10	3050.00	3139.50	3154.00
IRP	54,440.37	335.34	54,023.00	54,527.00	57,743.00
L	4.16	0.02	4.11	4.15	4.34
C	0.56	0.00	0.55	0.56	0.57
NC	30.97	0.49	30.00	31.00	36.00
S	0.90	0.00	0.89	0.90	0.91

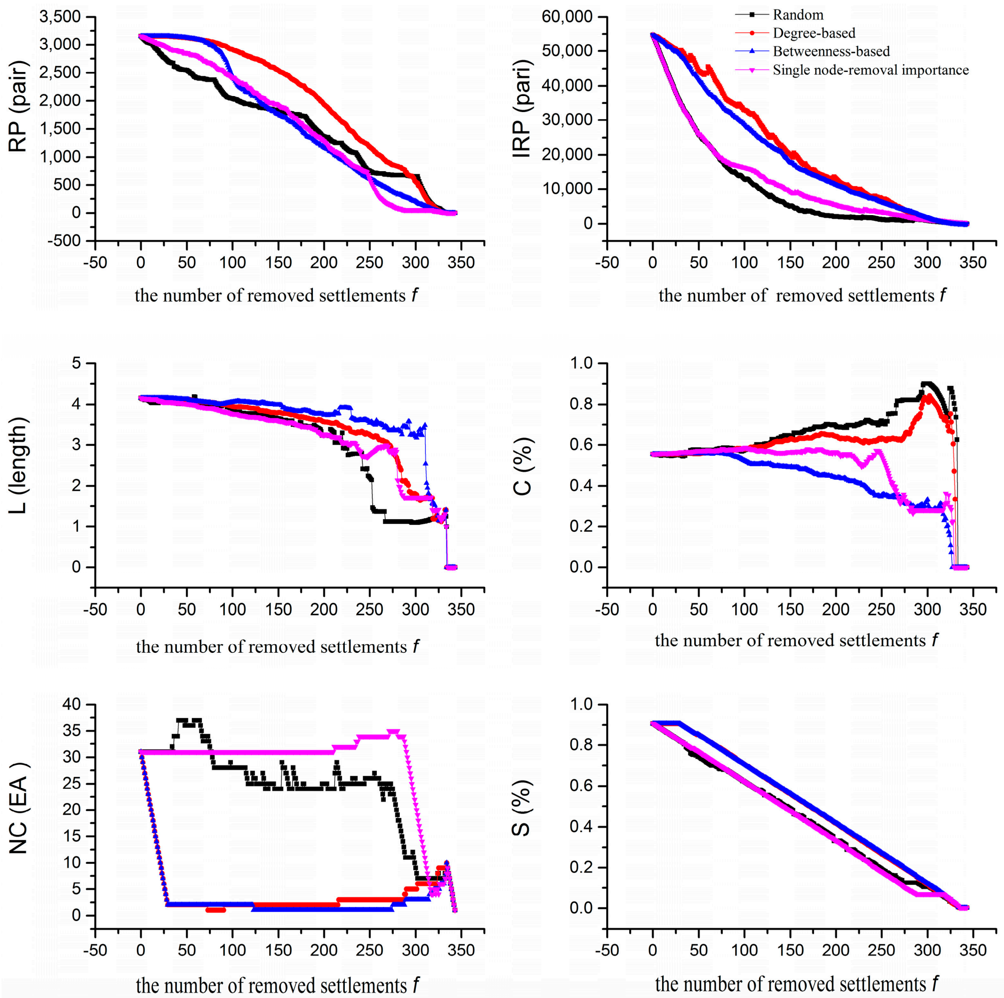

Figure 6. Typical behavior of socio-economic network under four successive node-removal attacks in function of the number of removed settlements f (RP: the number of directly reachable pairs, IRP: the number of indirectly reachable pairs, L: the average length of the shortest paths, C: the cluster coefficient, NC: the number of components, S: the relative size of the largest connected component).

Table 2. Relative change *r* of four metrics when half of settlements are removed.

Metrics	Random Strategy	Degree-Based Strategy	Betweenness-Based Strategy	Single Node-Removal Importance Strategy
RP	44%	26%	51%	47%
IRP	76%	41%	49%	71%
L	15.87%	11.58%	8.06%	16.38%
C	−21.95%	−13.99%	15.17%	−3.03%

Apart from the robustness of network, we also considered four traditional factors: (a) suitability (on a scale of 1–5); (b) dynamic change degree of source (see the study of Yang et al. [34] for details); (c) the level of hollowing (the proportion of unoccupied dwellings in all dwellings in each settlement); and (d) villagers' receptiveness of consolidation (the proportion of people who are willing to consolidate in all people in each settlement). The area of consolidated settlements in the target year (i.e., 2020) was obtained according to the population data from 1988 to 2014 derived from Daye Statistical Yearbook (see more in the study of Xuesong et al. [35]).

Through comprehensive consideration of the aforementioned five factors, we finally derived a classification map of source and sink (Figure 7b). To test the validity of the SRM, we also conducted a traditional classification (i.e., considering four traditional factors) (Figure 7a). Through contrastive analysis, we found that sources in traditional classification are mainly concentrated in the northern part of Chengui. Conversely, the spatial pattern of sources in optimized classification is dispersed and homogeneous. This divergence is caused by the discrepancy between traditional factors and socio-economic interactions. Specifically, although settlements in the north are relatively insufficient in suitability and villagers' receptiveness, they play an important role in the socio-economic village system (Figure 5).

We also conducted a detailed statistic comparison of two classifications (Table 3). The optimized classification shows higher potential in terms of maintaining the function of network system. The optimized strategy causes less harm to the entire village system than the traditional strategy especially for network connectivity and fragmentation (e.g., *RP* and *CN*). As for traditional factors, the optimized strategy also can obtain satisfied results except for villagers' receptiveness of consolidation (0.72 vs. 0.55). This exception indicates that the consolidation using optimized strategy faces challenges (e.g., the resistance of farmers). The local government may be able to effectively overcome these challenges (in terms of coordination of villagers' interests and sustainable development of the countryside).

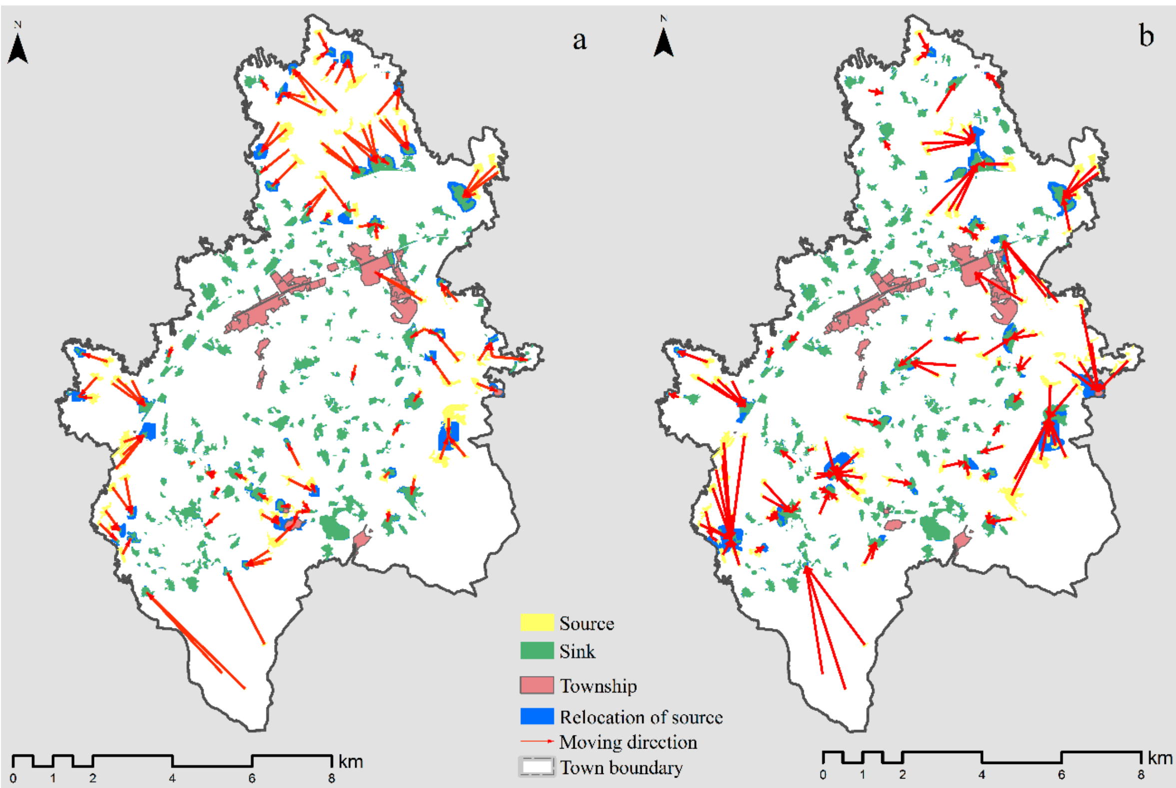

Figure 7. Classification and reorganization solutions in: (**a**) traditional; and (**b**) optimized scenarios.

Table 3. Comparison of two classification strategies.

Metrics	Traditional Scenario	Optimized Scenario	Positive (+)/ Negative(−)
Suitability of source	3.33	3.24	−
Suitability of sink	4.18	4.20	+
Dynamic change degree of source	0.01	0.01	−
The level of hollowing	0.88	0.87	+
Villagers' receptiveness of consolidation	0.72	0.55	+
RP after attack	1603	2368	+
L after attack	2.75	2.61	−
C after attack	0.63	0.68	+
CN after attack	132	119	−
S after attack	0.62	0.66	+

Note: The symbols (+) and (−) denote metric attributes. (+) represents a positive metric (higher score is better), and (−) represents a negative metric (lower score is better).

5.2. Implementation of GA to Relocate Consolidated Settlements

The typical parameter values of the GA are listed in Table 4. The population size was set to 100, and the number of generations to 300 (as the improvement of the best fitness value is stabilized after 200 generations). The crossover rate was set to 0.9, and mutation occurred in 5% of cases. For fully considering the socio-economic interactions and avoiding simple nearest reconstruction, d was set to 1000 m [36]. Based on the study of Yu [36], the scale coefficient α was set to 0.8. S, C, and L were all taken into consideration in optimized scenario ($w_S = 0.3, w_c = 0.3, w_l = 0.4$), whereas the traditional scenario only considered S and C with equal weights. The number of genes in optimized scenario was 137, whereas the length of chromosome in traditional scenario was 111.

Table 4. Parameter values of the genetic algorithm (GA).

Population Size	Iteration	Crossover Rate	Mutation Rate	d	α
100	300	0.9	0.05	1000	0.80

The solutions of spatial-territorial reorganization in different scenarios are presented in Figure 7. The moving distance (i.e., the distance from sources to relocated central settlements) of optimized solution is significantly longer than that of the traditional result (674.05 vs. 844.85). The moving direction of traditional scenario is more dispersed than that of the optimized solution. These differences suggest that the traditional strategy is a closer and dispersed integration, whereas the optimized strategy is wide-range and concentrated reconstruction. The optimized solution also shows more potential in terms of township development (i.e., the increased area of township, 15.37 hm^2 vs. 22.41 hm^2).

We also found two further advantages of the SRM through the detailed spatial comparison (Figure 8). First, the optimized scenario shows enhanced performance on the identification of administrative boundaries (Figure 8a,b). Specifically, the relocation of optimized strategy is under control of administrative system, whereas the traditional scenario seems to be under the control of distance. The reorganization solution of optimized scenario is always controlled within a village (the smallest administrative hierarchy in China), although it has longer moving distance. The cross-boundary relocation in traditional scenario may cause a separation between villagers even though they later live together. As reflected in our survey data, villagers in Wanjia rarely connected with others as they mostly came from other villages after natural disasters. This sense of belonging shared among villagers (located within a village) is formed by many factors (e.g., administrative and geographic context) in the long-term process of historical evolution. Second, the optimized scenario shows enhanced performance on the recognition of high-quality public service (e.g., shop, office, and school) (Figure 8c,d). The resettlement areas in optimized scenario are roughly twice as likely to own public services as the resettlement areas in traditional scenario (15.45% vs. 39.16%). Existing service facilities will largely reduce the relocation cost in a later period.

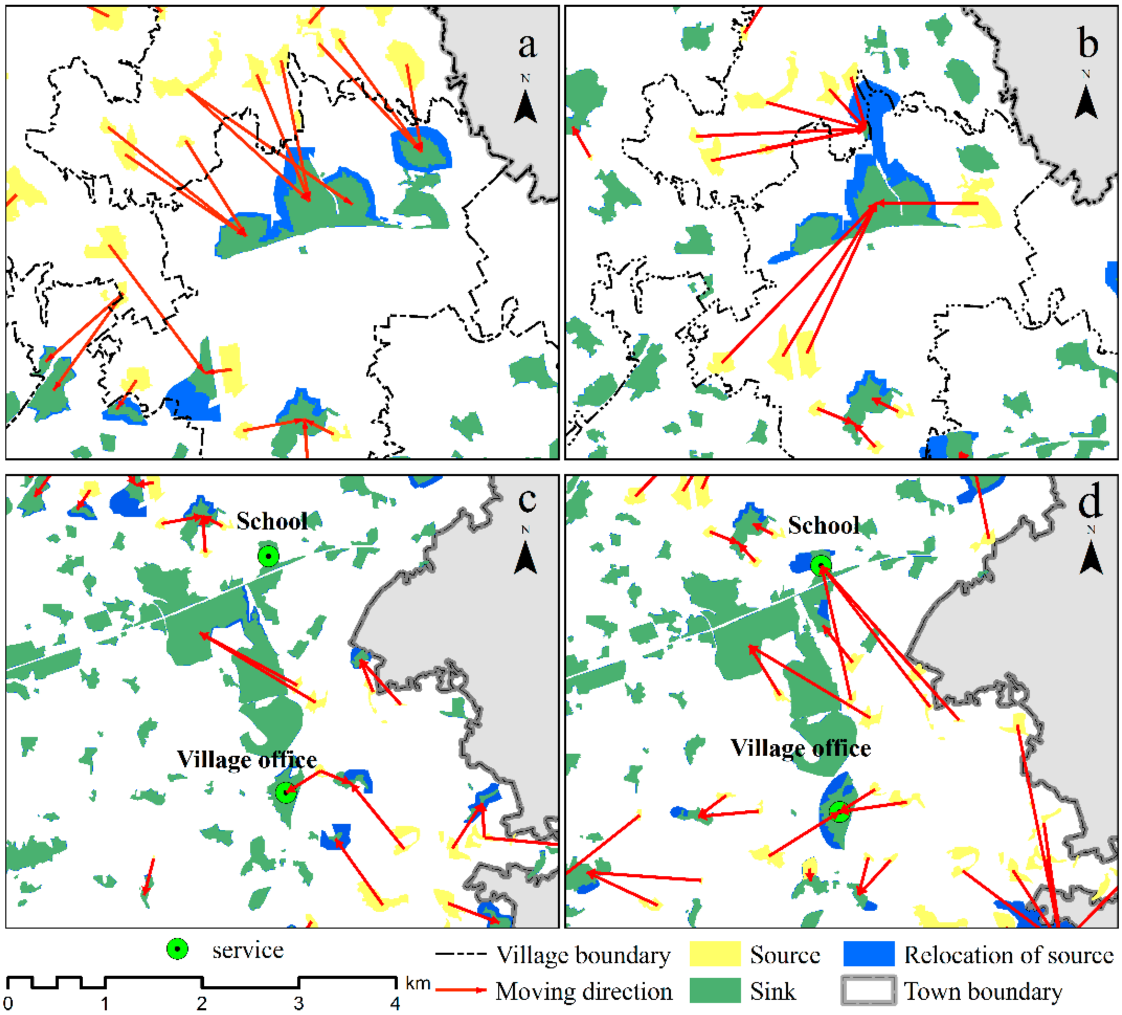

Figure 8. Different behaviors on: village boundary (**a**,**b**); and public services (**c**,**d**) (note: (**a**,**c**) represent the traditional scenario, while (**b**,**d**) represent the optimized scenario).

Moreover, the optimized and traditional solutions were compared with certain metrics to objectively prove the rationality of the SRM (Table 5). Significant improvements are observed in objective *L* and spatial metrics using SRM, whereas the traditional scenario has relatively better performance on objective *S* (4.12 vs. 4.06) and objective *C* (1.27 vs. 1.33). The optimized solution provides a local connectivity that is 40% higher than that of the traditional scenario (38.61 vs. 53.95). This efficiency demonstrates that the optimized solution has more demand for socio-economic interactions (working, visiting, shopping, and entertainment) from sources to sinks. The efficiency will be conducive for integration (e.g., lifestyle) between villagers in the resettlement area after consolidation. The receivable *L* value in traditional scenario is largely due to the propinquity effect—people who are located closer together in space show a higher probability of forming relationships [9,37–40]. Interactions among settlements are always bi-direction. In this context, we focused on the villagers' acceptability toward the central settlements because it is more crucial to the rural restructuring. Compared to traditional scenario, the *NP*, *PD*, *MPS*, *MPI*, and *MNN* of optimized solution have improved by 8.02%, 7.69%, 9.02%, 3.90%, and 2.23% respectively. In sum, the SRM helps to improve local connectivity and spatial pattern of rural settlements (e.g., bigger, more adjacent, and more concentrated) at the expense of a small reduction in suitability and compactness.

Table 5. Comparison of traditional and optimized scenarios.

Metrics	Traditional Scenario	Optimized Scenario	Positive (+)/Negative(−)
Weight strategy (*S*/*C*/*L*)	0.5/0.5/0	0.3/0.3/0.4	/
Object *S*	4.12	4.06	+
Object *C*	360.60	326.85	−
Object *C* (Mean, *C*/*H*)	1.27	1.33	−
Object *L*	38.61	53.95	+
The moving distance	674.05	844.85	−
NP	262	241	−
PD	2.34	2.16	−
MPS	4.88	5.32	+
MPI	94.28	97.96	+
MNN	104.25	101.93	−

Note: The symbols (+) and (−) denote metric attributes. (+) represents a positive metric (higher score is better), and (−) represents a negative metric (lower score is better).

The aforementioned scenarios have highlighted the significance of socio-economic interaction consideration in the process of reorganization. For different preferences/weights, the final solution may be understandably different [26]. To assess whether different combinations of the weight values would change the effectiveness of the model or solution significantly, we finally performed a preference analysis using different weight strategies. Herein, each objective preferred solution and solution with equal weights of object *S* and object *C* were taken as examples of the effectiveness of the SRM (Table 6). The classification result of source and sink came from the optimized scenario. The comparison table shows that all weight strategies demonstrate similar effectiveness in terms of object *S*, object *C*, as well as spatial pattern of rural settlements. This result implies the robustness of the SRM [26]. The obj-*S* and obj-*C* preferred solutions may not always reach the best scores with respect to their preferred single objectives, due to the small extent of study area and the characteristics of the objective *C* [26]. For obj-*L* preferred solution, the value of obj-*L* is much higher than the other solutions. Through contrastive analysis between four solutions and the aforementioned optimized scenario, we find that the optimized scenario has most balanced metric values. It is worth noting that it is arguably the best among these solutions because the aim of our reorganization method is to help planners or policy makers in finding suitable solutions according their preferences [26]. Regarding the metrics, we find that object *L* and moving distance are more sensitive to the weight strategy than suitability, object *C*, and landscape metrics. Meanwhile, we suppose that the pursuit of objective *C* and object *L* maybe face the problem of the increasing of moving distance.

Table 6. Attribute information associated with four objective preferred solutions.

Metrics	Equal Weights Solution	Obj-*S* Preferred Solution	Obj-*C* Preferred Solution	Obj-*L* Preferred Solution
Weight strategy (*S*/*C*/*L*)	0.5/0.5/0	1/0/0	0/1/0	0/0/1
Object *S*	4.06	4.06	4.06	4.06
Object *C*	325.38	329.27	324.75	326.81
Object *C* (Mean, *C*/*H*)	1.33	1.34	1.33	1.34
Object *L*	38.01	37.15	37.01	59.24
The moving distance	691.23	686.74	693.83	860.55
NP	239	240	239	238
PD	2.15	2.15	2.15	2.14
MPS	5.36	5.34	5.38	5.34
MPI	96.48	97.25	97.98	96.73
MNN	102.56	101.90	102.05	102.35

5.3. Relevant Policies/Practices

As the rural hollowing has become a major problem facing China's rural development, a series of policies, regulations, and practices have been introduced, such as "increasing vs. decreasing balance" land-use policy [4], "one family, one house" policy [3], and rural residential land consolidation and allocation (RRLCA) [6]. The aim of "increasing vs. decreasing balance" land-use policy is to achieve equilibrium in the supply of land through balancing the increases in urban construction land with decreases in rural construction land [4]. This policy has been developed at national and provincial levels to address the perceived rural problems [4]. The Constitutional policy of "one family, one house" has been proposed to address the problem of "outward expansion while inside hollowing" and "one family, more houses" [3]. RRLCA is an integrated approach to coordinate the numeric change of rural settlement and population [6]. It has been widely used and showed significant progress on the intensive utilization of rural residential land [6]. These policies are driven by the central state, represented by the Ministry of Land and Resources, and normally a top-down rural restructuring strategy [4]. They are insufficient for enrolling local actors into the planning and decision making [4]. The RRLCA can provide a gradually shift towards bottom-up endogenous development in the rural development strategies [6]. The SRM in this paper can serve as a local practice and participatory restructuring and planning method. With the help of related macro-policy platform and various RRLCA practices, the local practice of spatial-territorial reorganization can be advanced by local government [1]. Conversely, our local practice and SRM can provide an effective reorganization technique and optimization tool to realize these policies and practices. As such, accompanied by "top-down" policies/elements and platform provided by land consolidation, the local practice (SRM) and other "bottom-up" restructuring strategy are conducive to smoothly pushing forward spatial-territorial reorganization of rural settlement in China. The reorganization involves a diverse set of actors, such as local government, private enterprise, villagers, and other powerful actors, reflecting the complex and hybrid process of rural restructuring. The enrollment and motivation of these actors and participants in the process of consolidation could help to accelerate the reorganization/restructuring, safeguard farmers' benefits, and promote sustainable development in rural areas.

6. Conclusions and Future Work

This paper proposes a SRM of rural settlements to overcome the existing rural problems (e.g., scattered, massy, small, and hollowed). The proposed model is constructed based on graph theory and GA, and it involves two parts. In Part 1, we generate a series of experiments to investigate the network performance under numerous successive removals of settlement. Through this approach, we expect to find satisfied consolidated settlements with the least effect on the entire village system. In Part 2, GA model is repeatedly executed to optimize the objectives of suitability (S), compactness (C), and local connectivity (L) under the control of the constraints. The primary goal of this part is to scientifically resettle consolidated settlements into cities, nearby townships, or central settlements.

To verify the validity of the SRM, the proposed method has been applied in Chengui Town, Hubei Province. Two major findings are summarized as follows. First, removing settlements in order of node degree is the least efficient way to destroy the entire village system. Second, the proposed model can produce satisfactory solutions for spatial reorganization of rural settlements. The SRM helps to improve local connectivity and spatial pattern of rural settlements at the expense of a small reduction in suitability and compactness. The model also shows great potential in recognizing administrative boundary and high-quality public services.

The case study of Chengui is only a straightforward application of the SRM. It does not imply that the actual planning can be replaced. The optimization solution is merely a planning scenario based on different preferences. Moreover, the spatial-territorial reorganization of rural settlements is a complex systematic problem rather than a simple technological process. The reorganization may involve many aspects, such as the protection of basic farmland, industrial reorganization (off-farm employment), administrative reorganization, and land legal and managerial system [4]. The SRM may also be extended by including more objectives and/or constraints. Future research may analyze these perspectives in detail.

Acknowledgments: This work was supported by the National Key Research & Development Plan of China (Project Number: 2017YFB0503601).

Author Contributions: Yan Mao, Yanfang Liu, and Xuesong Kong conceived and designed the experiments; Yan Mao, Haofeng Wang, and Wei Tang performed the experiments; Yan Mao analyzed the data; Yanfang Liu, Haofeng Wang, Wei Tang, and Xuesong Kong contributed reagents/materials/analysis tools; Yan Mao wrote the paper.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Long, H.L. Land consolidation: An indispensable way of spatial restructuring in rural China. *J. Geogr. Sci.* **2014**, *24*, 211–225. [CrossRef]
2. Long, H.; Liu, Y.; Li, X.; Chen, Y. Building new countryside in China: A geographical perspective. *Land Use Policy* **2010**, *27*, 457–470. [CrossRef]
3. Liu, Y.; Fang, F.; Li, Y. Key issues of land use in China and implications for policy making. *Land Use Policy* **2014**, *40*, 6–12. [CrossRef]
4. Long, H.; Li, Y.; Liu, Y.; Woods, M.; Zou, J. Accelerated restructuring in rural China fueled by 'increasing vs. Decreasing balance' land-use policy for dealing with hollowed villages. *Land Use Policy* **2012**, *29*, 11–22. [CrossRef]
5. Liu, Y.; Liu, Y.; Chen, Y.; Long, H. The process and driving forces of rural hollowing in China under rapid urbanization. *J. Geogr. Sci.* **2010**, *20*, 876–888. [CrossRef]
6. Li, Y.R.; Liu, Y.S.; Long, H.L.; Cui, W.G. Community-based rural residential land consolidation and allocation can help to revitalize hollowed villages in traditional agricultural areas of China: Evidence from Dancheng county, Henan province. *Land Use Policy* **2014**, *39*, 188–198. [CrossRef]
7. Zhou, G.H.; He, Y.H.; Tang, C.L.; Yu, T.; Xiao, G.Z.; Zhong, T. Dynamic mechanism and present situation of rural settlement evolution in China. *J. Geogr. Sci.* **2013**, *23*, 513–524. [CrossRef]
8. Entwisle, B.; Faust, K.; Rindfuss, R.R.; Kaneda, T. Networks and contexts: Variation in the structure of social ties. *Am. J. Sociol.* **2007**, *112*, 1495–1533. [CrossRef]
9. Faust, K.; Entwisle, B.; Rindfuss, R.R.; Walsh, S.J.; Sawangdee, Y. Spatial arrangement of social and economic networks among villages in Nang rong district, Thailand. *Soc. Netw.* **1999**, *21*, 311–337. [CrossRef]
10. Zou, Y.; Liu, Y.; Kong, X.; Fan, D. Optimization of rural residential land based on weighted-voronoi diagram. *Geomat. Inf. Sci. Wuhan Univ.* **2012**, *37*, 560–563.
11. Duan, Y.; Lu, F. Robustness of city road networks at different granularities. *Phys. A Stat. Mech. Appl.* **2014**, *411*, 21–34. [CrossRef]
12. Boccaletti, S.; Latora, V.; Moreno, Y.; Chavez, M.; Hwang, D.U. Complex networks: Structure and dynamics. *Phys. Rep.* **2006**, *424*, 175–308. [CrossRef]
13. Albert, R.; Jeong, H.; Barabasi, A.L. Error and attack tolerance of complex networks. *Nature* **2000**, *406*, 378–382. [CrossRef] [PubMed]
14. Boldi, P.; Rosa, M.; Vigna, S. Robustness of social networks: Comparative results based on distance distributions. *Soc. Inform.* **2011**, *6984*, 8–21.
15. Li, X.; Yeh, A.G.O. Integration of genetic algorithms and gis for optimal location search. *Int. J. Geogr. Inf. Sci.* **2005**, *19*, 581–601. [CrossRef]
16. Liu, Y.; Tang, W.; He, J.; Liu, Y.; Ai, T.; Liu, D. A land-use spatial optimization model based on genetic optimization and game theory. *Comput. Environ. Urban Syst.* **2015**, *49*, 1–14. [CrossRef]
17. Stewart, T.J.; Janssen, R.; van Herwijnen, M. A genetic algorithm approach to multiobjective land use planning. *Comput. Oper. Res.* **2004**, *31*, 2293–2313. [CrossRef]
18. Fotakis, D.G.; Sidiropoulos, E.; Myronidis, D.; Ioannou, K. Spatial genetic algorithm for multi-objective forest planning. *For. Policy Econ.* **2012**, *21*, 12–19. [CrossRef]
19. Ducheyne, E.I.; De Wulf, R.R.; De Baets, B. A spatial approach to forest-management optimization: Linking gis and multiple objective genetic algorithms. *Int. J. Geogr. Inf. Sci.* **2006**, *20*, 917–928. [CrossRef]

20. Santé-Riveira, I.; Crecente-Maseda, R.; Miranda-Barrós, D. Gis-based planning support system for rural land-use allocation. *Comput. Electron. Agric.* **2008**, *63*, 257–273. [CrossRef]
21. Aerts, J.C.J.H.; Herwijnen, M.V.; Stewart, T.J. Using simulated annealing and spatial goal programming for solving a multi site land use allocation problem. *Lect. Notes Comput. Sci.* **2003**, *2632*, 448–463.
22. Liu, Y.L.; Wang, H.; Ji, Y.L.; Liu, Z.Q.; Zhao, X. Land use zoning at the county level based on a multi-objective particle swarm optimization algorithm: A case study from yicheng, China. *Int. J. Environ. Res. Public Health* **2012**, *9*, 2801–2826. [CrossRef] [PubMed]
23. Holland, J.H. *Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control, and Artificial Intelligence*; University of Michigan Press: Ann Arbor, MI, USA, 1975.
24. Goldberg, D.E. *Genetic Algorithms in Search, Optimization, and Machine Learning*; Addison-Wesley Longman Publishing: Boston, MA, USA, 1989.
25. Porta, J.; Parapar, J.; Doallo, R.; Rivera, F.F.; Santé, I.; Crecente, R. High performance genetic algorithm for land use planning. *Comput. Environ. Urban Syst.* **2013**, *37*, 45–58. [CrossRef]
26. Cao, K.; Huang, B.; Wang, S.; Lin, H. Sustainable land use optimization using boundary-based fast genetic algorithm. *Comput. Environ. Urban Syst.* **2012**, *36*, 257–269. [CrossRef]
27. Balling, R.J.; Taber, J.T.; Brown, M.R.; Day, K. Multi-objective urban planning using genetic algorithm. *J. Urban Plann. Dev.* **1999**, *125*, 86–99. [CrossRef]
28. Fotakis, D.; Sidiropoulos, E. Combined land-use and water allocation planning. *Annu. Oper. Res.* **2012**, *219*, 169–185. [CrossRef]
29. Boldi, P.; Rosa, M.; Vigna, S. Robustness of social and web graphs to node removal. *Soc. Netw. Anal. Min.* **2013**, *3*, 829–842. [CrossRef]
30. Pereira, M.; Segurado, P.; Neves, N. Using spatial network structure in landscape management and planning: A case study with pond turtles. *Landsc. Urban Plan.* **2011**, *100*, 67–76. [CrossRef]
31. Ding, C.R. Land policy reform in China: Assessment and prospects. *Land Use Policy* **2003**, *20*, 109–120. [CrossRef]
32. Brookes, C.J. A genetic algorithm for locating optimal sites on raster suitability maps. *Trans. GIS* **1997**, *2*, 201–212. [CrossRef]
33. Goldberg, D.E.; Deb, K. *A Comparative Analysis of Selection Schemes Used in Genetic Algorithms*; Morgan Kaufmann Publishers: San Mateo, CA, USA, 1991.
34. Yang, R.; Liu, Y.; Long, H.; Qiao, L. Spatio-temporal characteristics of rural settlements and land use in the bohai rim of China. *J. Geogr. Sci.* **2015**, *25*, 559–572. [CrossRef]
35. Kong, X.; Liu, Y.; Zou, Y.; Chen, Y. Calculation of land consolidation potential and optimization of rural residential areas based on households'willingness. *Trans. CSAE* **2010**, *26*, 296–301.
36. Yu, Z. Study on optimization of rural residential land based on weighted-voronoi diagram and calculation of consolidation potential. *Guangdong Land Sci.* **2013**, *12*, 43–48.
37. Adams, J.; Faust, K.; Lovasi, G.S. Capturing context: Integrating spatial and social network analyses. *Soc. Netw.* **2012**, *34*, 1–5. [CrossRef]
38. Hipp, J.R.; Faris, R.W.; Boessen, A. Measuring 'neighborhood': Constructing network neighborhoods. *Soc. Netw.* **2012**, *34*, 128–140. [CrossRef]
39. Festinger, L.; Schachter, S.; Back, K. *Social Pressures in Informal Groups*; Stanford University Press: Palo Alto, CA, USA, 1950.
40. Barthelemy, M. Spatial networks. *Phys. Rep.* **2011**, *499*, 1–101. [CrossRef]