Village-town System in Suburban Areas Based on Cellphone Signaling Mining and Network Hierarchy Structure Analysis

Ji-Zhe Zhou1, Quan-hua Hou1, Xiao-yang Fan1, Yang DU1

1 School of Architecture, Chang'an University, Xi'an, Shaanxi 710061, China
Corresponding author: Quan-hua Hou (e-mail: houquanhua@chd.edu.cn).

ABSTRACT Suburban villages are strongly influenced by cities. Understanding the relationship between urban and rural as well as the hierarchy of villages contributes to optimizing the structure of village-town, allocating resources effectively as well as coordinating the relationship between urban and rural areas. Taking the rural area of Qin and Han New City as the research object, this paper obtained the number of contact flows between urban and rural areas from cellphone signaling data mining and constructed a network with hierarchical system. The influence scope and hierarchy of city and village nodes in the network could be measured by conducting the primary linkage analysis. The effect intensity and radius of the central villages as well as the relationship between villages could be calculated by the multiple linkage analysis. Furthermore, according to the urbanization potential and the characteristic regions of cities, the regional village-town system could be finally determined. This study has shown that 89.4% of the villages were influenced by the surrounding cities strongly, and the influence range of the cities was 6.2-28.3km. There were 13 “Centrality” villages in this region, which was lean to city, traffic convenience and divisional centralized. Totally, 19 groups of villages should be merged and 3 groups should be developed together. Besides, two kinds of village-town systems were formed in this region, and the villages were classified into five categories for guidance. Moreover, the research results could not only provide a helpful method to determine the village-town system structure under the strong influence of cities but also offer a better way for resources allocation between urban and rural smartly and equally.

INDEX TERMS Suburban villages, Cellphone signaling data, Contact flow, Village-town system

I. INTRODUCTION
With frequent urban-rural contact, diverse travel destinations, and complex spatial characteristics [1], production and life in suburban villages are significantly influenced by economic production, cultural activities and public services of central cities [2]. According to the stage urbanization theory, urban development generally experiences four stages including urbanization, suburbanization, reverse urbanization and re-urbanization [3]. However, urbanization and suburbanization are coexisted in numerous suburban villages in China. The rapid urban development has attracted a large number of suburban populations into cities [4,5] and urbanization has been realized. Additionally, the outward development of cities is economical [6]. Suburban villages are affected by the excessive expansion of cities and face the dilemma of “passive urbanization” [7,8]. Under the policy guidance, “national new districts” have been established in China. As a kind of unnaturally formed urbanized areas, it includes the first two stages and reverse urbanization also occurs [9], leading to chaos in village-town system in suburban areas. The rapid urbanization has intensified the contradiction between urban and rural development in suburban villages [10,11]. There are not only emerging industrial bases, leisure and entertainment centers, commercial and trade logistics centers, but also abandoned land, idle public service facilities and population hollowing in suburban villages. Therefore, suburban village planning should balance the needs of different stakeholders [12], equalize the resource allocation as well as coordinate urban and rural development [13]. In the previous solutions, the United States focused on the coordination of workers and peasants and the joint development of urban and rural areas [14]. The United Kingdom adopted urban and regional planning to guide urban and rural development [15]. Netherlands and Germany formulated a rural renewal plan.

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/.
based on the land consolidation [16,17]. In the 1990s, China launched a village-city system plan to determine the village level at the regional level, lead rural development, and optimize the resource allocation. The complex structure [18] of suburban rural land leads to diverse purposes of population migration [19,20], and strong demands are raised for urban public services [21]. Through measuring regional city-village relationship and the village level system accurately, the village level structure and public facility allocation can be optimized, and rural development can be guided and controlled more accurately to achieve regional urban and rural sustainable development. The Qin and Han New City, located in the north of the Xi'an and Xianyang urban built-up areas, is affiliated to the Xixian National New District. Currently, approximately 90% of the Qin and Han New City is rural areas. Since rural areas are influenced by cities, it is the primary issue in rural planning to establish a suitable village system and rationally determine the control and development areas.

II. LITERATURE REVIEW

The earliest research on suburban village-town system is based on the evaluation of rural development “potential”, including land potential [22,23], building site selection [24], sustainability [25], tourism function evaluation [26,27], etc. Later, the “minimum cost” algorithm is included, such as the gravity model [28] and the service radius [29,30] to optimize the evaluation results. There are two major problems in the above-mentioned research which fails to consider the relevance and integrity of urban-rural relations and village-to-village relationships. First, the regional service and driving level of the “central village” selected for preferential development are much lower compared with the surrounding cities, and the population migrates to cities, causing rural resource allocation waste. Second, the guidance method of land use and population does not match the actual production and life, resulting in the lack or waste of land and facilities. In recent years, some scholars have studied the relationship and village level between villages and towns by setting a network hierarchy of tourism, society, and population [31,32,33]. The interaction between villages and towns is measured by the “Centrality” [34]. Based on the central place theory, “Centrality” is also taken as an indicator for measuring the central level and village-town system. For example, Hui measured the village-town system based on the contact degree, the contact direction and the contact trend between villages and towns [32]. The research emphasizes the regional integrity and relevance. However, 450 pieces of travel data collected from questionnaires can hardly measure the complex urban-rural relationship in suburban villages. Yang introduced commuting distance to measure the urban-rural relationship in suburban villages [34]. However, commuting distance cannot represent the actual urban-rural relationship, which should be measured based on time-space behaviors of human beings [35].

Therefore, to study the contact level between villages and surrounding areas, it is necessary to establish a regional contact network based on accurate time-space contact data. The village-town system can be measured through the “Centrality” and level of villages. Hierarchy is the basic feature of the contact network. Studying the network hierarchy helps to understand the relationship between network nodes and the structural characteristics of the network [36]. It provides a common way to divide the regional hierarchical structure according to the centrality of nodes in the mobile network by taking the flow intensity as the strength of the relationship between regions. For example, the port hierarchy structure is determined by analyzing the freight traffic in the shipping network [36,37]. The hierarchical structure of provincial capitals is determined by analyzing passenger and cargo flow volume in aviation and railway networks [38,39,40]. The urban hierarchy structure can be measured by analyzing the visitor flow volume in the population mobility network [41]. Combined with relevant research findings, the construction of network hierarchy at different levels raises different requirements for data sensitivity and accuracy, and more dynamic, continuous and accurate population flow information is required for suburban villages. The mobile phone signaling data can accurately measure the time-space behaviors of the suburban population [1,2,4,3,43]. Due to various advantages such as vector, large sample, full coverage, and continuous space-time attributes, it has been extensively used in regional urban system research and population flow characteristics analysis [44]. The existing research on the construction of contact networks based on mobile phone signaling data focuses on towns and cities, and suburban villages are seldom reported. Therefore, with the Qin and Han New City as an example, this paper constructs a network hierarchy by taking the mobile phone signaling data as the basic data to calculate the contact strength between the rural and surrounding areas. The centrality of villages is measured based on primary linkage analysis and multiple linkage analysis. Moreover, the influence strength and influence range of central villages are determined to predict rural development potential and urbanization characteristic areas. Finally, the regional village-town system is determined, hoping to provide reference and basis for rural facilities arrangement and overall urban and rural planning in suburban villages.

III. DATA AND METHODS

A. Research case and basic data

With a total area of 302.2 square kilometers, the Qin and Han New City is located among the urban built-up areas of Xi’an, Xianyang and Tonggang New City. It has jurisdiction over 142 administrative villages with a registered population of approximately 221,000. The planned new city is in the southwest of the area. At present, urban construction is underway. The present study aims to
explore the population mobility between each village and the surrounding cities and villages. With 142 villages and 3 urban built-up areas as nodes and the contact between objects as edges, the topological network diagram of regional population contact is constructed as shown in Figure 1.

The basic data is the 2G, 3G and 4G mobile phone signaling data of China Unicom collected from active and passive interaction between users and the base stations, including turning on/off, making calls, sending and receiving text messages and surfing the Internet, periodic location updates, location changes, conversion between 3G and 4G, etc. The recording time was 28 consecutive days in October 2018 (we also checked the October data by the recording of March). A total of 35 to 40 million records was collected. The records involved 88,000 permanent residents in suburban villages, accounting for 40% of the total registered population. In addition, the data also included the household registration data of 142 administrative villages, which were collected from the household registration offices of towns and streets. The road system data were collected from the Qin and Han New City New City Management Committee. The total area, cultivated land area, per capita income, infrastructure level, public service facility level, historical and cultural resources, and natural resources of the administrative villages was obtained from the basic material compilation in the village layout planning of Qin and Han New City in 2018. To check the regional population data, the actual population data of some regions were collected through the sample survey.

The multiple linkage analysis method was used to measure the influence intensity and the service radius of “Centrality” and analyze the relationship between rural nodes. Thirdly, the regional flow spatial features were obtained by analyzing the urbanization potential of rural nodes and urban characteristic regions. Finally, the “minimum cost” method based on public service facilities was introduced according to the regional space characteristics to check and improve the regional village-town system.

**FIGURE 1. Research object and its topology network diagram of population contact.**

**B. Analysis method and data processing**

1. Analysis framework

The overall research framework of “linkage features - spatial features - system construction” is shown in Figure 2. Firstly, based on the mobile phone signaling data, the primary linkage analysis method was used to preliminarily determine the influence range and the influence level of urban and rural nodes in the network hierarchy. Secondly, the multiple linkage analysis method was used to measure the influence intensity and the service radius of “Centrality” and analyze the relationship between rural nodes. Thirdly, the regional flow spatial features were obtained by analyzing the urbanization potential of rural nodes and urban characteristic regions. Finally, the “minimum cost” method based on public service facilities was introduced according to the regional space characteristics to check and improve the regional village-town system.

**FIGURE 2. Analysis framework.**

2. Linkage feature analysis method

(1) Cellphone signaling data processing

Urban and rural population mobility were determined in two steps: analyzing users’ stay points and the change of stay points. The changes of the stay points were mobility [1]. The stay points were located joint with the "Intelligent Footprint" company. Firstly, based on the location, power and coverage of the base stations, the area was divided into some square “base areas” (which may overlap). Generally, there were about 3 to 5 base stations in or around the “base area”. When the user accessed a new base station and continued to contact the next base station within the “base area”, or frequently cutting back to the original base station (signal overlap), and this state remained unchanged for half an hour, the user was identified as a “stay point”. The stay point was initially identified to be located in the “base area”. Secondly, according to the communication frequency and distance between the users in the “base area” and their neighboring base stations, the Location centroid algorithm was used to interpolate and calculate more accurate coordinates of the stay points [45]. The permanent stay point should be identified by gradually correcting the coordinate based on the data of next month. In general, the accuracy error of the coordinates of the stay point obtained by interpolation was 200-500 meters [1,45], and the corrected stop point error was less than 200 meters. There were two kinds of mobility: a user’s stay point changed in an hour and the new stop point was identified as the end point; When a user completed a travel but the stay point (within a threshold of 500m) had not changed in an hour, the farthest point he had reached was identified as the end point. If two different nodes recorded stay behaviors of the same user in one day, it was recorded as a flow between nodes (Figure 3). Due to the complex migration characteristics of suburban areas, multiple “flows” of the same user might be recorded in a day. The “flows” between the nodes in 28 days were counted in the present study.
Where, the 28-day city-village contact flow is \( X_{l,c} \), and the village-village contact flow is \( X_{l,l'} \). The formula is presented as follows:

\[
X_{l,c} = (C, V)_l c = \sum_{i=1}^{l} X_{l,c,i}
\]

\[
X_{l,l'} = (L, L')_l l' = \sum_{i=1}^{l} X_{l,l' i}
\]

Where: \( l \) and \( l' \) are the village node numbers (1, 1, 2, ...141), and \( c \) is the city node number (142, 143, 144); \( X_{l,c,i} \) is the contact flow between the village node \( l \) and the city node \( c \) on the \( i \) day; and \( X_{l,l' i} \) is the contact flow between the village nodes \( l \) and \( l' \) on the \( i \) day (\( i = 1, 2 ...28 \)); \((C, V)_l \) indicates that the starting point of the flow is the village and city. \((L, L')_l l' \) denotes that the starting point is the village \( l \) and \( l' \).

② Primary linkage analysis in urban and rural areas

The level of urban and rural nodes reflected their positions in spatial interaction. In order to study the regional urban-rural network structure, Nystuen et al proposed the primary linkage analysis (PLA) method [46]. The core idea was to determine the node level according to the flow direction of the maximum flow, that was, the dominant flow. As shown in Figure 4, if the dominant flow of a node flowed to a less important node, the node was called a dominant node. If the dominant flow of a node only flowed to more important nodes, and there was no dominant flow of other nodes flowed into, the node was called a dependent node. When the dominant flow of a node flowed to a more important node, and also had dominant flow of other less important nodes flowed into, the node was called a subordinate node. The importance degree of nodes refers to the relevant references [36,47]. Based on the external connection strength of urban and rural nodes, the importance degree of the nodes was calculated by the sum of formulas (1) and (2). In general, the largest or the top few dominant flows in suburban villages flowed to the surrounding cities. Therefore, this study should also include the number and direction of rural dominant flows in which urban flow was eliminated as the judgment standard for the “Centrality” of regional villages.

PLA could reflect the clustering and dominance of regional urban and rural nodes. In this study, the dominant flow of the rural node \( l \) and the surrounding area \( m \) (\( m = c, l' \)) was calculated. Additionally, the second and third largest dominant flows were also counted as the checking method. Statistical vector was \( \hat{q}_l \), and the formula was as follows:

\[
\hat{q}_l = [X_{l,1},X_{l,2},X_{l,3},X_{l,4},X_{l,5},X_{l,6}]
\]

Where \( X_{l,1} \) takes the village node \( l \) as the end point of the flow, contacts with the city node with the largest flow volume. \( X_{l,2} \) takes the village node \( l \) as the end point of the flow, and contacts with the village node with the largest flow volume.

(3) Multiple linkage analysis in urban and rural areas

Multiple linkage analysis (MLA) was the extended method based on the primary linkage analysis. Compared with the primary linkage flow analysis method, MLA could analyze the influence range and interaction of nodes. It was mainly determined by significant flows, that was, flows were greater than a given threshold. Initially, the flow volume flowing out of each node was arranged from large (\( X_l \)) to small (\( X_l \)), and the flow expectation set of the nodes was \( \{X_l\} \), \( j \in K \), \( K \) was the set of nodes in the network, and the calculation formula [36,48] was as follows:

Step 1: \( \tilde{X}_1 = \sum_{j=1}^{k} X_j, \tilde{X}_2 = \sum_{j=1}^{k} X_j, ..., \tilde{X}_k = 0 \)

Step 2: \( \tilde{X}_1 = \tilde{X}_2 = \sum_{j=1}^{k} X_j, \tilde{X}_2 = \sum_{j=1}^{k} X_j, ..., \tilde{X}_k = 0 \)

Step i: \( \tilde{X}_1 = \tilde{X}_2 = \sum_{j=1}^{k} X_j, \tilde{X}_i+1 = \sum_{j=1}^{k} X_j, ..., \tilde{X}_k = 0 \)

Step k: \( \tilde{X}_i = \tilde{X}_2 = \sum_{j=1}^{k} X_j, \sum_{j=1}^{k} X_j = 0 \)

Second, the degree of fit between the expected value and the true value of the flow volume was determined by the decision coefficient \( r^2 \), and the calculation formula was shown in Equation (4). If the decision coefficient \( r^2 \) in step j was the largest, the flows before j (include j) were all significant flows.

\[
r^2 = 1 - \frac{\sum_{j=1}^{k} (X_j - \tilde{X}_j)^2}{\sum_{j=1}^{k} (X_j - \tilde{X})^2}
\]

According to the number and flow direction of the significant flow, the importance degree and influence range of the village nodes could be judged. Table 1 presented the meanings of the nodes corresponding to different significant flow relationships, which could be taken as a basis for determining the relationship between villages and towns. For example, if there was a certain type of village in the region was the significant flow direction of multiple villages, it had “Centrality” and regional service function, and could be taken as a candidate central village. If two adjacent villages A and B had significant rural flows to each other, they could coordinate development and even merged. If the significant flow of
village A only flowed to B, B should be the incorporation destination of A (migration).

3. Spatial feature analysis method

The analysis of flow spatial characteristics involved urbanization potential and development trend. Urbanization level was analyzed by employing the method of studying information flow in American metropolitan areas proposed by James O. Wheeler [49]. The villages with urbanization potential were classified by the natural breakpoint method [41]. Rural urbanization potential \( C_l \) was calculated as follows:

\[
C_l = \ln \frac{x_{1l}}{x_{Tl}}
\]  

(5)

Table 1. Flow directions and meanings of significant flows.

<table>
<thead>
<tr>
<th>Description</th>
<th>The significant flow of A only flows to B</th>
<th>The significant flows of both A and B flow to each other</th>
<th>The significant flow of A flows to two or more nodes</th>
<th>The significant flows of two or more nodes flow to A</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spatial structure</td>
<td>A (\rightarrow) B</td>
<td>A (\leftarrow) B</td>
<td>A (\ldots) C_i</td>
<td>A (\ldots) C_i</td>
</tr>
<tr>
<td>Meaning</td>
<td>Village A has strong dependence on B</td>
<td>When adjacent, village A and B are complementary or independent</td>
<td>Potential competition exists between the target villages in village A</td>
<td>Village A has centrality</td>
</tr>
</tbody>
</table>

4. Construction method of village-town system

In the suburban village-town system, the “central village” had the functions of connecting urban functions, driving regional development, and serving the surrounding villages as well as had a concentrated advantageous resource allocation [50,51]. To establish a village-town system, it was necessary to highlight the leading role of the “central village” and control the cost while maximizing the utility. The “minimum cost” algorithm should be introduced to improve the layout results. In suburban villages, the villagers were primarily concerned with the equalization of educational facilities in urban and rural areas [52]. Primary schools, as the major public facilities with high frequency of use by villagers and also the public facilities with the farthest daily service radius in rural areas [53,54], often served as the service radius threshold of the “minimum cost” algorithm [55,56]. Field research results indicated that the commuting time that could be endured by local residents was 15–25 minutes, and the commuting mode was motor vehicles (the local motor vehicle penetration rate was as high as 91%, and its speed was 15km/h). The “minimum cost” algorithm based on the primary school service radius was completed on the GIS platform to establish regional residential areas and road network models. Referring to the service radius index of the primary school in the adjacent Jingyang County [53], high-quality primary schools would be established in villages with “Centrality” feature as the candidate central village. The radius was calculated from 3750m to 6250m, and increased by 50m each time. The calculate could be cut off until the full coverage of urban and rural public services as well as the serviced population of each “central village” was over 4000 (the minimum indicators of population that need for primary school). The radius for cut off was the “minimum cost” of public service. According to the result, the number of the candidate central villages would be increased or decreased to determine the regional village-town system of Qin and Han New City.

IV. ANALYSIS RESULTS

A. Regional linkage features

1. Regional linkage analysis results

With the regional cities and villages nodes as the end points, the contact flows between villages and cities were calculated in pairs and the contact flows in 28 days were summarized as shown in Figure 5. The flows from villages to cities accounted for 73.4% of the total urban-rural linkages, indicating that cities strongly affected the development of villages. The contact of Xi’an, Xianyang, and the northern city with villages accounted for 53.2%, 29.3% and 17.4%, respectively. Based on the attraction scale of the village nodes in the network, the importance of the nodes was determined. The natural breakpoint method [41] was employed to divide the 142 villages into five important degrees. Through checking according to the basic data, the rural areas with high importance were the main employment, tourism and transportation hubs in Qin and Han New City. The rural classification and main
functional areas were presented in Figure 6.

FIGURE 6. Main functional areas and rural node importance degree.

2. Primary linkage analysis results

The dominant flow of villages and cities in Qin and Han New City were analyzed to determine the influence degree and influence range of cities on villages, as presented in Figure 7 and Table 2. Among the 142 administrative villages in Qin and Han New City, the targets of rural dominant flows of 127 villages were cities, and 89.4% villages were within the strong influence range of cities. Through analyzing the commuting distances of the farthest dominant flow from villages to cities, the maximum influence radiuses of the regional urban centers in Xi’an, Xianyang and northern urban areas were 28.3km, 15.2km and 6.2km, respectively. The regional urban-rural contact volume was positively correlated with the maximum impact radius and the size of the surrounding cities. According to the primary flow analysis results, the dominant linkage flows between 40 villages and the surrounding cities were much greater than those with other rural areas. They were generally located at urban fringe or the border areas, having the initial functions and characteristics of cities. In addition, the villages which frequent contact were also surrounding other cities, and the contact between villages reflects trans-regional characteristics.

Table 2 Influence degree and influence range of surrounding cities on villages.

<table>
<thead>
<tr>
<th>City</th>
<th>Villages influenced by dominant flow</th>
<th>Villages influenced by the second dominant flow</th>
<th>Villages influenced by the third dominant flow</th>
<th>Maximum influence radius</th>
</tr>
</thead>
<tbody>
<tr>
<td>Xi’an</td>
<td>51</td>
<td>46</td>
<td>18</td>
<td>28.3km</td>
</tr>
<tr>
<td>Xianyang</td>
<td>51</td>
<td>22</td>
<td>14</td>
<td>15.2km</td>
</tr>
<tr>
<td>Northern city</td>
<td>25</td>
<td>34</td>
<td>25</td>
<td>6.2km</td>
</tr>
</tbody>
</table>

The dominant flow of regional villages was analyzed to determine the centrality and rank of villages at the network level. According to the analysis results in Figure 8, there were 12 dominant nodes, 42 subdominant nodes and 88 dependent nodes in the region. In the analysis of dominant nodes, two dominant villages near the Xi’an built-up area on the southeast side of the region had a high degree of contact with cities. However, the contact with the surrounding villages was weak without dominant flow from other sub-villages. It could be preliminarily determined that they had a tendency of integrating to cities. There were three such villages in Qin and Han New City. The remaining nine dominant villages were the end points of regional rural flows, and seven of them were judged to have “Centrality” due to the large number of dominant flows or regional centralized flow direction. Since some subdominant villages were also the concentrated flow direction of regional dominant flows, the rural areas to which the second dominant and third dominant flows flowed should be counted to determine their level in the network system.

FIGURE 8. Primary linkage analysis results based on rural dominant
flows.

A total of 13 villages had “Centrality” in this region (Table 3). The analysis results demonstrated that “Centrality” villages were generally located in the flow path from secondary villages to the city which biased toward the city as well as regional centers far from cities. “Centrality” villages had strong regional concentration characteristics and were the ending point of regional people flow. The village-village dominant flow line in the network hierarchy fitted well with the main road network, and 76.9% “Centrality” villages were located around important intersections.

Table 3. Village “Centrality” analysis results.

<table>
<thead>
<tr>
<th>Node type</th>
<th>Name of the village</th>
<th>Dominant flow</th>
<th>Second dominant flow</th>
<th>Third dominant flow</th>
<th>Judgment reason</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dominant</td>
<td>Nanbeishangzhao</td>
<td>10</td>
<td>0</td>
<td>9</td>
<td>Large number of dominant flows</td>
</tr>
<tr>
<td></td>
<td>Liwei</td>
<td>7</td>
<td>6</td>
<td>1</td>
<td>Large number of dominant flows</td>
</tr>
<tr>
<td></td>
<td>Houying</td>
<td>4</td>
<td>1</td>
<td>2</td>
<td>Regional core</td>
</tr>
<tr>
<td></td>
<td>Dongfeng</td>
<td>6</td>
<td>3</td>
<td>1</td>
<td>Large number of dominant flows</td>
</tr>
<tr>
<td></td>
<td>Weichengwan</td>
<td>4</td>
<td>1</td>
<td>2</td>
<td>Regional core</td>
</tr>
<tr>
<td></td>
<td>Yaodian</td>
<td>5</td>
<td>2</td>
<td>4</td>
<td>Large number of dominant flows</td>
</tr>
<tr>
<td></td>
<td>Lingzhao</td>
<td>13</td>
<td>14</td>
<td>5</td>
<td>Large number of top two dominant flows</td>
</tr>
<tr>
<td></td>
<td>Yanjiagou</td>
<td>5</td>
<td>6</td>
<td>4</td>
<td>Large number of dominant flows</td>
</tr>
<tr>
<td></td>
<td>Baimiao</td>
<td>6</td>
<td>7</td>
<td>2</td>
<td>Large number of top two dominant flows</td>
</tr>
<tr>
<td></td>
<td>Yiwei</td>
<td>8</td>
<td>2</td>
<td>4</td>
<td>Large number of dominant flows</td>
</tr>
<tr>
<td></td>
<td>Dawang</td>
<td>6</td>
<td>3</td>
<td>5</td>
<td>Large number of third dominant flows</td>
</tr>
<tr>
<td></td>
<td>Shuangzhao</td>
<td>3</td>
<td>3</td>
<td>7</td>
<td>Large number of dominant flows</td>
</tr>
<tr>
<td></td>
<td>Guxian</td>
<td>2</td>
<td>4</td>
<td>5</td>
<td>Regional core</td>
</tr>
</tbody>
</table>

3. Multiple linkage analysis results

Primary flow analysis could simplify the regional urban-rural linkage network structure, yet lead to the loss of information about the important edge, which flow was second only to the dominant flow but significantly superior other flow [36]. The multiple linkage analysis method could analyze the flow which exceeds a certain threshold as a significant flow to solve the above problem effectively. The central village in the village-town system should have the ability to drive the development of the surrounding area, and provide high-quality public services. Consequently, the more the significant flow of other villages flowed to the central village, the more important it was in the network level, should develop with superior resources. For the villages in Qin and Han New City, since the inflow of cities was much larger than that of the surrounding villages, the significant flows of cities and villages should be calculated and summarized separately.

The multiple linkage analysis results were shown in Figure 9, which checked the judgement of the “Centrality” village by the dominant flow. Based on the number of significant flows, the regional network hierarchy was divided into five levels. The “Centrality” village was located in the first two levels, and the number of influent significant flows was greater than 12. The “Centrality” village only had regional centrality. It was the direction of regional significant flow and also had the maximum number of flows compared with the surrounding areas. The rural areas in the urban fringe had more significant flows. Some “Non-centrality” villages even had more significant flows than the “Centrality” villages in suburbs and this phenomenon was attributed to the attraction of cities to regional population.

FIGURE 9. Multiple linkage analysis results.

Based on the direction of significant flow, the main affected scope (villages) of the “Centrality” villages was...
further determined. If a village had two significant flows flowing to different “Centrality” villages, the one with larger significant flow was the village to which it belonged. There were 13 centralized areas and 1 non-centralized areas, which can be found Figure 10. “Centrality” villages had a large difference in influence ability. The largest “Centrality” village was Lingzhao, which was located in the regional center and had 22 villages. The smallest “Centrality” village was Shuangzhao, located in the northwest of the region and had 4 villages. The “Centrality” villages in the region were generally not located in the geometric center, yet in the regional contact hub or middle and lower reached areas of urban-rural linkages. Although some villages were located in remote areas or enclaves, they make frequent communication with the “Centrality” villages, and the commuting time was equivalent to that of other villages. Combined with the actual commuting distance between rural settlements, the influence intensity and the service radius of “Centrality” villages could be obtained. As shown in Table 4, there were 4 “Centrality” villages in the urban fringe area, which had 27-37 significant flows and affected 6-22 villages. The average maximum influence radius was 6,750 meters. There were 9 non-urban fringe “Centrality” villages, which had 13-23 significant flows and affected 4-11 villages. The average maximum influence radius was 5400 meters. The “Centrality” villages in the urban fringe had higher significant flow intensity and better regional influence ability. The average significant flow doubled that of the “Centrality” villages on the non-urban fringe. The number of affected villages was 1.7 times, and the influence radius was 1.25 times. However, the difference in the average influence radius was small. Based on the flow of significant flows between villages, their relationship could be further judged. As shown in Figure 10, 19 groups of villages had the merger potential. They were either located in the outer suburbs and driven by urbanization (farther one immigrated to the nearer one) or in the urban fringe, which led to the extremely complex urban-rural linkages and the urbanization trend. There were 3 groups of rural areas with the synergistic development potential. They either had similar sizes, resources and traffic conditions, or shared industrial, commercial facilities and infrastructure.

<table>
<thead>
<tr>
<th>Location</th>
<th>Name of village</th>
<th>Number of significant flows</th>
<th>Number of influence villages</th>
<th>Maximum influence radius</th>
<th>Average influence radius</th>
</tr>
</thead>
<tbody>
<tr>
<td>Urban fringe</td>
<td>Nanbeishangzhao</td>
<td>37</td>
<td>14</td>
<td>7400m</td>
<td>2800m</td>
</tr>
<tr>
<td></td>
<td>Dawang</td>
<td>27</td>
<td>10</td>
<td>5600m</td>
<td>3900m</td>
</tr>
<tr>
<td></td>
<td>Yanjiagou</td>
<td>27</td>
<td>6</td>
<td>4800m</td>
<td>4100m</td>
</tr>
<tr>
<td></td>
<td>Lingzhao</td>
<td>36</td>
<td>22</td>
<td>9200m</td>
<td>5100m</td>
</tr>
<tr>
<td>Non-urban fringe</td>
<td>Guxian</td>
<td>14</td>
<td>5</td>
<td>4000m</td>
<td>3100m</td>
</tr>
<tr>
<td></td>
<td>Liuwei</td>
<td>15</td>
<td>11</td>
<td>6300m</td>
<td>3500m</td>
</tr>
<tr>
<td></td>
<td>Houying</td>
<td>23</td>
<td>8</td>
<td>5500m</td>
<td>3300m</td>
</tr>
<tr>
<td></td>
<td>Shuangzhao</td>
<td>17</td>
<td>4</td>
<td>3300m</td>
<td>2500m</td>
</tr>
<tr>
<td></td>
<td>Weichengwan</td>
<td>13</td>
<td>6</td>
<td>6100m</td>
<td>4100m</td>
</tr>
<tr>
<td></td>
<td>Yaodian</td>
<td>13</td>
<td>8</td>
<td>7200m</td>
<td>4000m</td>
</tr>
<tr>
<td></td>
<td>Yiwei</td>
<td>17</td>
<td>11</td>
<td>6100m</td>
<td>3400m</td>
</tr>
<tr>
<td></td>
<td>Baimiao</td>
<td>16</td>
<td>7</td>
<td>5800m</td>
<td>5200m</td>
</tr>
<tr>
<td></td>
<td>Dongfeng</td>
<td>13</td>
<td>9</td>
<td>4800m</td>
<td>3200m</td>
</tr>
</tbody>
</table>

B. Flow spatial features

1. Urbanization potential analysis results

The villages with the urbanization potential were shown in Figure 11, which was divided into six levels. The average urbanization potential of the villages was 0.68, and the urban-rural linkages in villages on both sides of the value were balanced. The villages in the first two levels had outstanding urbanization potential. Totally, they were 40 such villages, accounting for 28.6% of the total. The areas with high urbanization potential were mainly concentrated in the urban fringe of Xi'an, Xianyang and the northern city; Qin and Han New City was carrying out urban construction areas and its northern affected areas. Among the 17 villages with the urbanization potential at the first level, 14 villages had carried out urbanization construction within its scope of administration, and 8 rural settlements have completed the demolition and resettlement work. The 23 villages with the urbanization potential at the second level were mainly concentrated in or around large industrial parks, large commercial facilities, leisure and holiday attractions, ruins & museums, and tourist villages.
The urban areas and urbanized central village regions. The study results showed that the primary public service facilities in the region were located in the southeast of the region. The study results demonstrated that 93.7% of the serviced villages were in the “Centrality” village service area, and could both meet the needs of service radius and commuted preferences.

The result of village-town system was shown in Table 5. There were two types of village-town systems and five types of rural guidance in the Qin and Han New City. The three-level system of “urbanized central village - urbanized village - village” was located on the side of the cities in the region, covering seven urbanization villages and their range of influence. The main public services in the region were provided by the urbanized central villages. Urban facilities allocation standards were adopted. All of public service facilities in the region were shared by villages and cities. Urbanized villages should be urbanized recently, and the general village should develop towards cities gradually. The two-level system of “central village - village” was located on the west and north sides of the region, covering seven central villages and their range of influence. Although basic public services were provided by the central villages, important public services were provided by cities. The general village should gradually develop towards regional central village.

FIGURE 11. Villages with urbanization potential.

2. Urban characteristic regions analysis results

Based on the urban-rural contact intensity, with the average flow impact distance of 2700m as the search radius, and 100*100m as the smallest unit, the kernel density difference analysis was performed in GIS [40]. The flow space aggregation intensity was presented in Figure 12. The regional flow space was characterized by hierarchy, multi-center, urban-rural differences, forming three first-level cores of Lingzhao, Weichengwan and Houpai, and three second-level cores of Dawang, Nanbeishangzhao and Yanjiagou. There were two urban development corridors from Xianyang to the northern city and along the new urban built-up area of Qin and Han. The urban areas and rural areas in the region had obvious differences. Based on the analysis results of urbanization potential and related urban and rural planning, by referring to the multi-core urbanization development model in suburbs [57,58], the urbanization characteristic regions of Qin and Han New City were determined, which can be found in Figure 12. A total of 55 villages was covered. To be specific, 7 of the 13 “central” villages were located in urbanized areas, and the affected areas were divided into three levels of “urbanized central village - urbanized village - village”. Other 6 villages were located in non-urbanized areas, and the affected areas have a “central village - village” two-level village-town system structure.

FIGURE 12. Regional flow space aggregation intensity and urbanization feature regions.

C. Results of village-town system

According to the above analysis, primary schools could be established in 13 “Centrality” villages. The maximum service radius was 6000 meters. As shown in Figure 13, the service range of the originally selected centrality village could not satisfy the needs of full coverage of public service, and it was necessary to add a central village. Based on the primary linkage and multiple linkage analysis results, the central village added should be Tongren, located in the southeast of the region. The study results demonstrated that 93.7% of the serviced villages were in the “Centrality” village service area, and could both meet the needs of service radius and commuted preferences.


FIGURE 14. Development guidelines for rural classification under two types of village-town systems.
Development guidelines for rural classification under two types of village-town systems were shown in Figure 14. The detailed construction guidelines were as follows:

1. Urbanized central village. They were taken as the regions that relieve the population and function of surrounding cities, developing business and service industries, and should drive regional development. Public facilities such as hospitals, vocational schools, primary and secondary schools, cultural and sports centers, large department stores, and nursing homes would be provided to achieve the equalization of public services in urban and rural areas.

2. Urbanized villages. They should promote the construction of urban communities and building the community public service facilities. Encourage the land transfer, gradually introduce secondary and tertiary industries; should provide skill training and guide the urbanization of rural residents. A certain proportion of low-cost buildings would be established recently in local areas to solve the housing needs of migrant workers.

3. Central villages. They should optimize the industrial structure, improve the industry quality, provide preferential condition, guide the gathering of surrounding villages, and build new rural communities. Besides, kindergartens and primary schools, the old-age service stations, health service centers, supermarkets, farmers’ markets and other public facilities should be deployed according to rural standards to serve the surrounding villages.

4. Reserved villages. They should optimize industrial structure, improve farming methods, upgrade basic public services and infrastructures, and improve the rural living environment which could reflect unique characteristics. Kindergartens, elderly activity rooms, health stations and outdoor event venues should be established.

5. Demolished villages. Villages gradually merge into the target village (Figure 10). Residents could work in cities and towns through the land transfer. Additionally, they could also farm in other villages, through land replacement. Housing construction should be restricted, and public facilities should meet the basic needs of daily life.

### V. DISCUSSIONS AND CONCLUSIONS

#### A. Discussions
2. Application of mobile phone signaling data in village-town system planning

Based on the mobile phone signaling data, the population property and time-space distribution information in suburban villages could be obtained comprehensively, dynamically and accurately \([62,63,64]\), and the contact level could be accurately measured. However, there still remained a problem of recognition accuracy. This study had some errors in taking the changes of stay point as the judgment standard of “flow”. Although the flow of a user per day was within a limited range, there were 1-24 flows theoretically. Therefore, it was challenging to estimate the actual population flow from contact flow. The error was due to the use of desensitized cell phone signaling data. For the sake of privacy protection, the desensitized data could not reflect personal travel trajectories. Relevant research indirectly identified the flow of people with certain characteristics by determining the appearance conditions and intervals of users at different moments, identifying resident populations \([1]\), permanent residences \([1,65]\), workplaces \([65,66]\), shopping populations \([67]\), travel modes \([68]\), etc. For the research on mobile phone signaling data of Qin and Han New City, our team had accurately identified the population composition, commuting characteristics, and urban-rural linkage characteristics of the region. Due to the complex composition of the regional population, the large number of external populations, the variety of travel modes as well as unfixed time for work, shopping and rest, the relevant human flow research would have large errors in local applications. However, by analyzing the flow amount, for the construction of regional network hierarchy, the study of regional village-town system could still obtain similar results compared with by actual population flow \([41]\).

3. Application of primary linkage and multiple linkage analysis in village-town system planning

The primary linkage analysis aimed to study the maximum flow and the flow direction of nodes in the network. It could not only effectively determine the intensity and direction of urban and rural population flow, the influence radius and the influence degree of cities, but also predict the levels and development trends of villages and cities. This study optimized the recording of multiple flows by Li (2016) \([36]\). The results demonstrated that recording the inflow of significant flows from other villages could better reflect the importance of rural nodes in the region than recording significant flows in the received flow. It could effectively avoid the neglected the needs of vulnerable villages caused by overpopulation and resource-rich villages which led the proportion of significant flows, reflected the idea of “equality” in planning, and better guided the equalization of regional resources. Since the two analysis methods had different emphases, the results would be slightly different. Therefore, they should be jointly used in research. In the existing village-town system, the main role of “central villages and towns” was to provide goods and daily services to the surrounding villages as well as to guide the development of the region. It was essentially evolved from the “Central geographic model” proposed by W. Christaller. The existing research on village-town system planning gradually developed from the “Central geographic model” into a “Life circle model”. After determining the regional core, the public service facilities would be arranged step by step in the inner and outer circles \([56,69]\). The existing research methods of the village-town system based on “potential evaluation” or “minimum cost” selected the center of the living circle, divided the circle and determined the reasonable service radius, which was a continuation of the central geographic model. As shown in Fig. 16, the central villages and towns selected by the traditional method were biased toward the geometric center of the region. However, in the suburban areas of megacities or national new districts, the urban-village relationship and the village-village relationship had been beyond the scope of the central geographic model. Central villages were generally not in the geometric center of the region, and were located in an area with dense urban-rural flow. Public service facilities might not satisfy the principle of proximity, which might be distributed in remote enclaves or even cities based on road links. In such village-town system, the analysis method of primary linkage and multiple linkage would exert an important role.

4. Application of the research in urban and rural public service allocation

According to the survey results, it could be found that the vacancy rate of rural public service facilities was extremely high. For example, only 18 of the 44 primary schools in the region have reached the used capacity. Field
interviews showed that rural households tended to use primary schools in further urban fringe areas. The allocation of public facilities in suburban villages depended largely on villagers’ preferences rather than commuting distances. Based on usage preferences, the equipped public service facilities helped to the shrewd used of urban and rural public service facilities. Regional public services were mainly provided in urban fringe, which had the following advantages: ① Rural residents enjoy urban public services with minimal commuting distance; ② through bidirectional flow of urban and rural population [1], rural residents could enjoy urban public services during commuting hours, like sending children to school on the way to working; ③ rural areas had less serious traffic congestion. Although the theoretical commuting distance was greater, the actual commuting time was equivalent to that of urban areas. In this study, through analyzing linkage features and spatial features, and identifying urbanized central villages which provided regional services were also located in the fringe of urbanized areas.

B. Conclusions

Based on mobile phone signaling data, the research adopted theory of network hierarchy and the analysis method in transportation engineering field to measure the centrality of village and its affected range through primary linkage analysis and multiple linkage analysis. Moreover, the village-town system was established according to urbanization potential and urban characteristic areas by taking Qin and Han New City as an empirical case. The results had shown that village-city contacts account for 73.4% of total contacts, and 89.4% villages were strongly influenced by cities. The influence radiuses of the surrounding Xi’an, Xianyang and the northern city were 28.3km, 15.2km and 6.2km, respectively. Villages in urban fringe began to have urban functions and characteristics, and external contact reflected cross-regional characteristics. The 13 “Centrality” villages were generally located in the flow path from secondary villages to the city which biased toward the city, or regional centers far from cities. They had the characteristics of convenient transportation and regional concentration. The average significant flow of “Centrality” villages on the urban fringe doubled the “Centrality” villages on the non-urban fringe, the number of influenced villages was 1.7 times, and the maximum affect radius was 1.25 times. However, they had small differences in the average influence radius. There were 19 groups of villages with merger potential, and 3 groups of villages with synergistic development potential. The rural areas with high regional urbanization potential were mainly concentrated on the fringe of three cities and the areas influenced by the new city. There were 55 villages which preliminary had urbanization characteristics. Finally, in the Qin and Han New City, a village-town system with “urbanized central village - urbanized village - village” and “central village - village” was established, which would be guided by five types of villages.

Compared with the previous research, firstly, the research methods of the existing village-town system were extended. The original intention of the village-town system planning was to centrally allocate superior resources and guided rural development smartly. Under the strong influence of cities, people’s flow behaviors could reflect the complicated urban-rural relationship in the suburban areas. The disclosure of villagers’ commuting preferences made up for the waste of resource allocation caused by the potential evaluation based on objective factors. Secondly, the object of network level was further expanded. The network level was analyzed from the perspectives of international aviation and the port system in the field of transportation engineering, which was gradually applied to study city-town systems and even village-town systems in the field of urban planning. Moreover, a more complete information analysis system of goods flow, information flow and people flow could be formed. Finally, this study had also solved the problem of village-town system planning under the strong influence of cities, made reasonable speculations on the urbanization characteristic areas, and determined the development guidelines for each village. Moreover, the research results could provide reference and basis for studying spatial structure optimization in suburban areas, public service facilities layout, urban and rural development, rural policy formulation, urban boundary speculation and demarcation.

Our team will further introduce urban and rural ecological flow, information flow and material flow into the network, and analyze the robustness and vulnerability of urban-rural networks based on complex network model. Moreover, by strengthening the ability of urban-rural network system to resist urban uncertain disturbances, the resilience of regional villages will be improved under the influence of severe urbanization. However, the research on the influence mechanism of urban and rural flow prediction, rural space characteristics, contact flow on villagers’ livelihood also requires field research and analysis of villages in areas with different characteristics.

ACKNOWLEDGEMENTS

This research was funded by the National Natural Science Foundation of China (Grant No. 51978058), Fundamental Research Funds (Social Science) for the Central Universities of China (Grant No. 300102419631) and China Scholarship Council (Grant No. 201906560052).

REFERENCES


This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2019.2939504, IEEE Access
Quan-Hua Hou received the Ph.D. degrees in transportation planning and management from Chang’an University. He is a professor and deputy dean of School of Architecture, Chang’an University. He has achieved numerous scientific researches of urban and rural planning, and published more than 20 papers which indexed by SCI and EI. He has led or participated in 10 research projects at the national or provincial level. In recent years, he has focused on the rural planning, and has rich theoretical and practical basis.

Email: houquanhua@chd.edu.cn

Xiao-Yang Fan is currently a graduate student majoring in urban and rural planning in School of Architecture, Chang’an University. The research direction mainly in school is rural planning theory and method.

Email: fxy960307@163.com

Yang Du she is currently a graduate student majoring in urban and rural planning in School of Architecture, Chang’an University. The research direction mainly in school is rural planning theory and method.

Email: Dduuy@163.com