applied sciences

MDPI

Article

Identification of Critical Road Links Based on Static and Dynamic Features Fusion

Yi Li and Min Huang *

School of Intelligent Systems Engineering, Sun Yat-sen University, Shenzhen 510275, China; liyi89@mail2.sysu.edu.cn
* Correspondence: huangm7@mail.sysu.edu.cn

Abstract: Traffic congestion is a significant challenge in modern cities, leading to economic losses, environmental pollution, and inconvenience for the public. Identifying critical road links in a city can assist urban traffic management in developing effective management strategies, preserving the efficiency of critical road links, and ensuring the smooth operation of urban transportation systems. However, the existing road link importance evaluation metrics mostly rely on complex network metrics and traffic metrics, which may lead to biased results. In this paper, we propose a critical road link identification framework based on the fusion of dynamic and static features. First, we propose a directed dual topological traffic network model that considers the subjectivity of road links, traffic circulation characteristics, and time-varying characteristics, which addresses the limitations of existing traffic network topology construction. Subsequently, we employ a novel graph representation learning network to learn the road link node low-dimensional embeddings. Finally, we utilize clustering algorithms to cluster each road link node and evaluate critical road links using the average importance evaluation indicator of different categories. The results of comparison experiments using real-world data demonstrate the clear superiority and effectiveness of our proposed method. Specifically, our method is able to achieve a reduction in traffic network efficiency of 70–75% when less than 25% of the road links are removed. In contrast, the other baseline methods only achieve a reduction of 50–70% when removing the same proportion of road links. These findings highlight the significant advantages of our approach in identifying the critical links.

Keywords: urban traffic network; critical links; network representation learning; intelligent transportation system; data mining

check for updates

Citation: Li, Y.; Huang, M. Identification of Critical Road Links Based on Static and Dynamic Features Fusion. *Appl. Sci.* **2023**, *13*, 5994. https://doi.org/10.3390/app13105994

Academic Editor: Xinlin Huang

Received: 20 April 2023
Revised: 9 May 2023
Accepted: 10 May 2023
Published: 13 May 2023

1. Introduction

The phenomenon of urbanization has led to a surge in travel demand in urban areas, which poses a significant challenge to urban traffic. However, the infrastructure development in these areas is constrained by the limitations of urban land, resulting in an inability to meet the changing traffic demand. Traffic congestion is emerging as a major urban problem, posing various challenges in terms of environmental pollution, time wastage, and reduced productivity. The urban traffic network is a time-varying, directed, and weighted network that should account for both its topological structure and travel characteristics. Given the vast size of the urban traffic network, it is not feasible to manage each road link during traffic control. Due to resource and funding limitations, we can only manage a limited number of urban road links. Therefore, it is crucial to determine which road links should be controlled to maintain the normal operation of urban traffic. Additionally, in the event of natural disasters, it is necessary to prioritize which road links to restore or control first to improve the operational efficiency of the urban traffic network. By evaluating the importance of urban road links, we can identify critical road links in the city and prioritize their control and maintenance to ensure the smooth operation of the urban traffic network.

The urban road network serves as the foundation of the urban transportation system and is a crucial component of transportation infrastructure. Research [1] indicates that a

cascading failure phenomenon exists in urban road networks, whereby the failure of a few links can lead to the failure of other links and ultimately affect the efficiency of surrounding local networks. Critical links refer to those links that, when failed due to traffic congestion or accidents, have a significant impact on network operating efficiency. Therefore, it is imperative to identify such critical links and devise management strategies to address the issue of traffic congestion and ensure transportation efficiency. Evaluating the importance of links in the traffic network is significant to solving this problem. By analyzing urban road network topology and traffic data to identify bottlenecks and congestion hotspots, urban traffic management can focus their efforts on the most critical links in the network. This targeted approach can improve traffic flow and reduce congestion, ultimately leading to a more efficient and sustainable transportation system.

In the traffic system, traffic travel has been widely linked to the urban spatial structure, and the connection has been extended to the field of urban science theory by many researchers. Some researchers have explored the urban operation mechanism and analyzed the causes of traffic congestion [2] by using large-scale spatio-temporal data. Their analysis involves identifying urban activity hotspot areas [3], urban spatial density [4], and urban spatial structure [5]. Meanwhile, some researchers have combined complex network theory and traffic travel characteristics to explain traffic phenomena. They use quantitative indicators commonly used in complex networks and graph theory, such as degree, centrality, and network efficiency, and analyze the relationship between these indicators and traffic travel to measure the importance of links in the traffic network [6,7]. Such research efforts can provide valuable insights into the urban traffic phenomenon and pave the way for the development of effective traffic management strategies.

In this study, we propose a traffic link importance evaluation method based on combining road network topology with traffic demand and traffic state through network representation learning. Firstly, we establish a directed dual traffic network model using the dual mapping method based on a directed road network, where road links are represented as nodes and the connection relationships between them as edges. Secondly, we employ a novel representation learning method to obtain the embedding vectors of the link nodes in the network and utilize machine learning techniques to cluster the acquired embedding vectors. Finally, we define importance evaluation indicators by combining traffic demand and traffic state and analyze the importance of different categories of link nodes after clustering in order to identify the critical link nodes in the network.

The main contributions of this paper are as follows:

- We propose a method for constructing a traffic network with a directed dual topology. The method emphasizes the main position of urban road links in the traffic system. It also reflects the circulation characteristics between urban road links and the time-varying characteristics of the importance of urban road links.
- We propose a novel representation learning method. The method combines urban traffic network structure and urban traffic demand to jointly control the random walk procedure and learn low-dimensional representations of road link nodes using skip-gram models.
- We design an efficient method for evaluating the importance of urban road links. Combining the embedding vectors of road link nodes, the clustering algorithm and the road link importance evaluation indicators are used to evaluate the road link importance, which can accurately identify the important road links in the city.
- To validate the effectiveness of the proposed method, we designed and compared experiments combining real-world data to analyze the performance of each method. The experimental results validate the effectiveness and superiority of the proposed methods.

2. Literature Review

In recent years, the evaluation of road link importance has been mostly based on the construction of road network topology using the primal method for undirected road networks. Two primary research approaches are commonly used in this field: quantitative evaluation using importance indicators and assessing changes in network efficiency following road link failures.

In terms of quantitative evaluation using importance indicators, some researchers combine complex network theory indicators and traffic information for quantitative analysis. Girvan et al. [8] introduced the concept of edge betweenness, which is based on betweenness centrality. This concept suggests that road links with higher edge betweenness have a greater transmission capacity and a more significant role in the network. Wang et al. [9] proposed a node importance discrimination method based on local features, which considers the importance of neighboring nodes in addition to the node itself. Tu et al. [10] introduced the minimum cut frequency vector to construct road link importance evaluation indicators for road link vulnerability identification evaluation between OD pairs. Wang et al. [11] used the Fuzzy C-means (FCM) clustering algorithm to analyze the importance of road links by combining betweenness, PageRank, and traffic flow. Yi-Run et al. [12] combined node degree and node local similarity to evaluate the importance of road links. Su et al. [13] proposed a model for evaluating the importance of road links under different time delays by considering the spatio-temporal correlation between multi-order neighboring road links. Chen et al. [14] proposed a method to identify the importance of road links in the "effective impact area" by considering that the failure of a road link mainly affects the surrounding neighboring areas and calculating the efficiency of the regional network in this way to reduce computation time. When dealing with smaller transportation networks, importance indicators can provide satisfactory results for quantitative evaluation. However, as the size of the network increases, the calculation of indicators based on complex network theory can become a time-consuming process and the evaluation results may not accurately reflect the actual traffic conditions.

In terms of assessing variation in network efficiency following road link failures. Some researchers have employed traffic assignment theory to quantify the impact of road link failures on traffic network efficiency. The critical road links are then identified by comparing the change in traffic network operation efficiency before and after the failure of these links. Scott et al. [15] proposed the network robustness index (NRI) measure to identify the importance of road links by combining network capacity, road network traffic flow, and topology. Jenelius et al. [16] used the user equilibrium (UE) model to reassign network traffic flow after road link failure and identified the importance of road links by comparing the cost of travel on the network after the failure of different road links. Sun et al. [17] combined link failure probability and comprehensive indicators of network efficiency and total travel time of the road network after traffic reassignment to evaluate the importance of road links. Zhang et al. [18] developed a model for link importance assessment that considers various factors such as road network structure, traffic demand, travel behavior characteristics, and multi-link failures by using the travel time of all users in the system as a measure of traffic network performance.

The current research on link importance evaluation in traffic networks faces two main challenges. Firstly, most road network modeling is based on undirected graphs using the primal method [19], which maps intersections as nodes and connecting links as edges. The approach ignores the circulation relationship between road links and the time-varying characteristic of road link importance. In urban transportation, the circulation relationship between road links is a fundamental aspect of the traffic system, as it characterizes the traffic sources and destinations of each link. Moreover, the importance of each direction of a road link varies over time, particularly during peak hours such as morning and evening rush periods. Consequently, road links may have vastly different levels of significance depending on the time of day, necessitating distinct control strategies. Secondly, extracting indicators for evaluating the importance of road links mainly relies on complex network theory and

traffic indicators. However, complex network theory indicators require manual effort and are time-consuming. As the size of the network increases, the feature matrix becomes high-dimensional and sparse, making it unsuitable for further analysis. Additionally, prior research typically utilizes a single traffic indicator without considering the combination of multiple traffic characteristics. The work presented in this paper aims to address and overcome these deficiencies.

3. Methodology

3.1. Directed Dual Traffic Network Constructing

Most existing research on traffic network modeling typically represents intersections as nodes and the connections between intersections as edges. However, in the field of traffic, the road link plays a crucial role, serving as the primary location of traffic accidents and traffic control measures in traffic networks. To address this limitation, we construct a directed traffic road network model using the dual method [20], which preserves the connection and circulation relationships within the traffic network. Specifically, in this method, links are mapped as nodes, and the connection relationships between different links are represented as edges, while the circulation relationship of traffic flow between links is captured as connections. To capture the traffic demand and state in the city, we calculated the traffic flow and average speed of road links within a specific time interval. The directed dual traffic network is defined as follows:

$$G = (V, E, S_T, F_T) \tag{1}$$

where $V = \{v_1, v_2, \ldots, v_n\}$ is the set of link nodes which means the road links, and n is the number of road links. $E = \{(v_i, v_j), v_i, v_j \in V\}$ is the set of edges in the network. $S_T = \{s_1, s_2, \ldots, s_n\}$ is the set of average speeds of link nodes in a specific time interval T. $F_T = \{f_1, f_2, \ldots, f_n\}$ is the set of traffic flow of link nodes in a specific time interval T. T is a constant, such as 30 min.

Compared to the traffic network with the primal method, the directed dual traffic network emphasizes the significant role of links in the traffic system. The approach enhances the clarity of the topological relationship between different links. In the directed dual traffic network model, the static and dynamic characteristics of road links can be captured. The structure of the traffic network displays the connection relationship between different links, which depicts the static characteristics of road links. Meanwhile, the traffic flow and average speed of the road links reflect the dynamic characteristics of the traffic network. Where the traffic flow characterizes the traffic demand, and the average speed characterizes the traffic state. It is crucial to note that a small traffic flow during congested times may not necessarily indicate low traffic demand.

Based on the connection relationships between road links in the directed road network, the original topology of the road network can be well represented as the directed dual traffic network. Figure 1 illustrates the original directed traffic network topology and its corresponding directed dual traffic network topology. In the figure, letters A-H represent different intersections, while numbers 1–7 represent different road links. The direction of the road links indicates the direction of traffic flow. For instance, in the original traffic network, road link 1 and road link 4 have a flow connection, whereas road link 1 and road link 2 do not. Consequently, in the directed dual traffic network, road link node 1 is connected to road link node 4, but not to road link node 2.

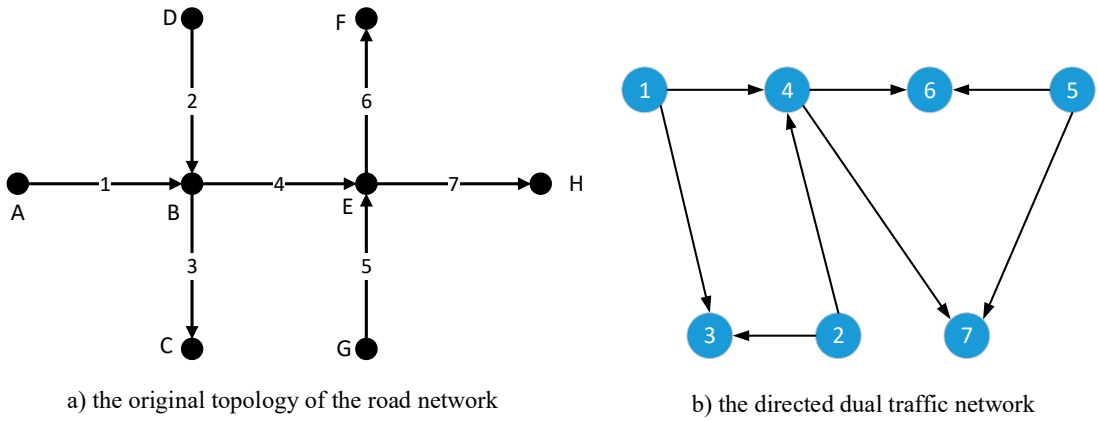

Figure 1. Example of directed dual traffic network.

3.2. Traffic Link Representation Learning

After constructing the directed dual traffic network G, various indicators can be calculated to evaluate the importance of links. However, the extraction of importance indicators, such as adjacency matrix, degree, betweenness centrality, and network efficiency, from complex networks is a time-consuming process. Additionally, as the size of the network increases, the feature matrix obtained from these methods exhibits high dimensionality and sparsity, rendering clustering-based evaluations of link importance barely satisfactory.

Graph representation learning (GRL) can address these limitations well, which is a technique that enables the extraction of effective features from graph information. One model that is widely used for feature extraction in GRL is Node2vec. Based on the Deepwalk model [21], Node2vec [22] integrates walk bias, which incorporates depth-first search (DFS) and breadth-first search (BFS) balance coefficients to consider both local and global features of nodes. Node2vec, introduced in 2016, takes the network structure as input and outputs the embedding vector of each node. Its objective is to map the node information in the network to a low-dimensional, continuous, and dense feature space. The embedding vectors both preserve the neighborhood features of the original node and facilitate subsequent data mining.

In this paper, we propose a novel representation learning model named TraLink2vec, which extends the Node2vec method by integrating both static and dynamic features in the traffic network. TraLink2vec inherits the walk bias of Node2vec for the static road network structure from the perspectives of BFS and DFS, while also incorporating the dynamic travel features of the traffic system to construct dynamic travel bias from the perspective of traffic demand. To sample the directed dual traffic network G, TrLink2vec simulates a random walk process of fixed length l for r times. Given a current node c_{i-1}, the next node c_i is generated based on the following distribution:

$$P(c_i = x \mid c_{i-1} = v) = \begin{cases} \frac{\pi_{vx}}{Z} & \text{if } (v, x) \in E \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

where π_{vx} is the unnormalized transition probability between nodes v and x, and Z is the normalization constant.

In order to balance the dynamic and static features in the traffic network, we use a second-order random walk procedure to guide the walk. Suppose the current random walk passes through the edge (t, v) to reach v. We set the unnormalized transition probability $\pi_{vx} = \pi(x \mid t, v)$ as follows:

$$\pi(x \mid t, v) = \begin{cases} \alpha \times \frac{1}{p} \times w_{vx} + \beta \times \frac{\varepsilon_{flow\_x}}{\varepsilon_{speed\_x}} & \text{if } d_{tx} = 0 \\ \alpha \times w_{vx} + \beta \times \frac{\varepsilon_{flow\_x}}{\varepsilon_{speed\_x}} & \text{if } d_{tx} = 1 \\ \alpha \times \frac{1}{q} \times w_{vx} + \beta \times \frac{\varepsilon_{flow\_x}}{\varepsilon_{speed\_x}} & \text{if } d_{tx} = 2 \end{cases} \tag{3}$$

$$\varepsilon_{flow\_x} = \frac{flow_x}{flow_v} \tag{4}$$

$$\varepsilon_{speed\_x} = \frac{speed_{design}}{speed_x} \tag{5}$$

where $\pi(x \mid t, v)$ is the unnormalized transition probability of the next link node x through the edge (t, v). d_{tx} is the shortest path distance between link node t and link node x. w_{vx} is the weight between link node v and link node x, which means the length between the midpoints of the two link nodes. Parameter p is the return parameter, which controls the probability of returning to the previous link node. Parameter q is an in-out parameter, which controls the probability of returning to the previous link node or exploring far-away link nodes. Parameter α is the control weight of static features and parameter β is the control weight of dynamic features. ε_{flow_x} is the traffic demand coefficient, $flow_x$ is the traffic flow of link node x and $flow_v$ is the traffic flow of link node v. ε_{speed_x} is traffic state coefficient, $speed_x$ is the average speed of link node x and $speed_{design}$ denotes the design speed of link node x.

As shown in Figure 2, the traffic flow and traffic speed on the nodes represent the dynamic attributes of the link nodes. The connection relationship between nodes represents the static attributes of the link nodes. In terms of static features that affect the walk, the label values on the edges characterize the control parameters of the next link node x passing link node v. Where the control parameter for returning to the previous link node t is $\frac{1}{p}$. The control parameter for exploring further nodes such as link node x_2 and x_3 is $\frac{1}{q}$. The control parameter for exploring other link nodes such as link node x_1 is 1. In terms of dynamic features that affect the walk, link nodes with higher traffic flow and lower traffic speed are found to have a higher transition probability. For instance, in the case of link nodes x_2 and x_3, the transition probability from v to x_2 is higher than from v to x_3. Conversely, for link nodes t and x_3, the transition probability from v to t is higher than from v to x_3.

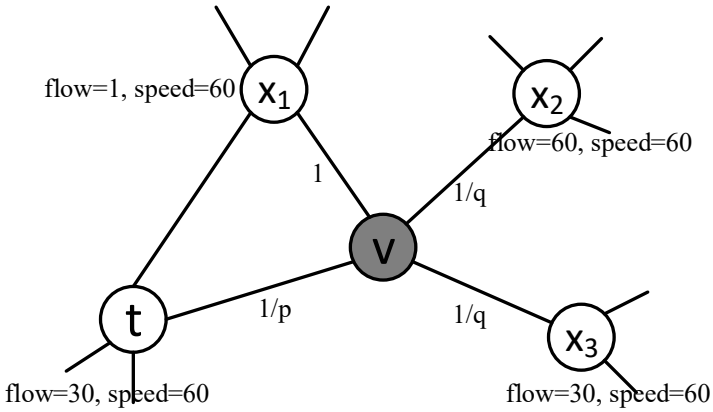

Figure 2. Illustration of random walk process in TraLink2vec.

Therefore, with $v_i (i = 1, 2, \ldots, n)$ as the starting link node, the walk sequence $W_{v_i} = \left\{ w^1_{v_i}, w^2_{v_i}, \ldots, w^l_{v_i} \right\}$ can be obtained by simulating a random walk through the transition probability function, where $w^j_{v_i} (j = 1, 2, \ldots, l)$ denotes the jth link node passed in the walk procedure with v_i as the starting link node, l denotes the length of the walk. Repeat r times for each node in the traffic network, and finally obtain r walk sequences $W = \left\{ W_{v_i,1}, W_{v_i,2}, \ldots, W_{v_i,r} \right\}$ of length l for each node, where $W_{v_i,k}$ denotes the kth repeat of the walk sequence with v_i as the starting link node. Finally, the walk sequences obtained from each node in the traffic network are utilized as input, and the skip-gram model, which is commonly used in natural language processing (NLP), is employed for training. This

allows for the mapping of road links with dynamic and static attributes into d-dimensional vectors.

3.3. Traffic Link Importance Evaluation

In previous studies, clustering was often used to evaluate the importance of road links after obtaining some link importance characteristics. The importance of road links is typically classified into three categories: critical, normal, and unimportant [11]. Following this concept, the embedding vectors obtained by TraLink2vec are clustered in this paper. Generally speaking, the importance of a road link is positively related to its traffic demand, with higher traffic demand indicating greater importance. Traffic demand is directly reflected by the traffic flow of the road link, and higher traffic flow indicates higher traffic demand. Additionally, the traffic state of the road link can also characterize its traffic demand. When a road link is congested, its traffic flow may be small, but this does not necessarily mean that its traffic demand is also small. Therefore, the importance evaluation indicator of road link is defined as follows:

$$I_v = flow_v \times \frac{speed_{design}}{speed_v} \tag{6}$$

where $flow_v$ is the traffic flow of link node v, $speed_v$ is the average speed of link node v, $speed_{design}$ is the design speed of link node v. This indicator represents the ratio of traffic demand to traffic state in a road link. The larger I_v means that the traffic demand on the road link is greater as well as the road link is more important.

After computing the importance evaluation indicator of each link node, the average importance of each clustering category is then calculated based on the clustering results, which serve as the evaluation index of road link importance. The traffic link importance evaluation process is demonstrated in Algorithm 1.

Algorithm 1 Traffic Link Importance Evaluation

Input: Link node, V; Link length, W; Traffic flow, F_T; Average speed of road, S_T; Link design speed, S_{design}; Time interval, T;
Output: Nodes label and important rating, IR

```
G = Constructing Directed Dual Traffic Network (V, E, S_T, F_T);
Calculating embedding vectors of link nodes E(V) = TraLink2vec(G);
Clustering embedding vectors of traffic networks C = clustering(E(V));
Calculating the importance evaluation indicator i = index(EV)
Initializing the importance of link, IR = list();
for v in V do
    for c in C do
        if cluster results of v ∈ c then
            v.ir = calculating the average importance of link of all V ∈ c;
        end if
    end for
end for
for v in V do
    Calculating importance rank result r = sort by v.ir
    IR. append (v, r)
end for
```

In this paper, we utilized the K-means clustering algorithm [23] to cluster the embedding vectors due to its simplicity and efficiency. The algorithm initiates by randomly selecting k samples as cluster centers and then computes the distance between each sample and the cluster center. Subsequently, each sample is assigned to the closest cluster center, and the samples assigned to each center form a cluster. After all samples are allocated, the centroid within each cluster is recalculated and reassigned until it converges to a stationary position. This clustering method is well-suited for datasets with clusters of arbitrary shapes.

4. Experimental Setup

4.1. Dataset Description

In this paper, the central area of Huangpu District in Shanghai was selected as the study area, and the basic data included two data sources. The first one was the GIS map of the urban area of Huangpu District in Shanghai, which included road direction, road level, design speed, number of lanes, and other relevant information. The second one was the automatic vehicle identification (AVI) data in Shanghai during January, which contained the AVI device number, the desensitized license plate of the vehicle, vehicle type, passing time, driving direction, lane number, and other relevant information. We reconstructed the travel trajectories of all vehicles based on the travel time estimation method and mapped individual travel behavior to the road network for calculation. This process enabled us to obtain the traffic flow and average speed of each road link during a specific time interval.

We constructed a directed dual traffic network in the central area of Huangpu District using the data sources described above, as illustrated in Figure 3. This network comprised 277 road link nodes and 769 edges. By integrating the AVI data, we were able to obtain the traffic flow and average speed of each road link during a specific time interval.

a) road network in the central area of Huangpu District b) the directed dual traffic network

Figure 3. Directed dual traffic network construction.

4.2. Evaluation Metrics

(1) In this study, we adopt the Calinski–Harabaz Index (CHI) [24] to evaluate the performance of the clustering results and to identify the optimal parameters for the representation learning algorithm. The CHI is calculated as follows:

$$CHI = \frac{tr(B_k)(m-k)}{tr(W_k)(k-1)} \tag{7}$$

where m is the number of link nodes, k is the number of clusters, B_k is the covariance matrix between different clusters, W_k is the covariance matrix between data within clusters, $tr(\cdot)$ is the traces of matrix. This indicator is the ratio of all inter-class distances to intra-class distances. A higher value of CHI indicates that the intra-classes are more tightly packed, and the inter-classes are more widely separated, which leads to a better clustering effect.

(2) The transportation network is a typical open and complex system. The failure of a few links can have a cascading effect, ultimately impacting the efficiency of the surrounding local networks. Therefore, we consider using the variation of network efficiency [25] to evaluate the performance of the proposed method in this study.

In the urban transportation system, the traffic demand of each road link is different. Generally speaking, road links with higher traffic demand have a greater impact on the surrounding road links. Based on this principle, we propose a method for calculating traffic network efficiency that considers urban travel demand. We define the traffic network efficiency as follows:

$$e_i = \frac{1}{(n-1)} \sum_{i \neq j} \frac{1}{d_{ij}} \times I_i \tag{8}$$

$$E(G) = \frac{1}{n} \sum_{i \in V} e_i \tag{9}$$

where e_i is the link node efficiency of link node v_i, d_{ij} is the shortest path distance between link node v_i and link node v_j, I_i is the importance evaluation indicator of the link node v_i. $E(G)$ is the overall traffic network efficiency, which is the average link node efficiency of all link nodes in the traffic network. n is the number of link nodes. The link node efficiency characterizes the importance of nodes in urban traffic networks, with a larger e_i indicating a greater importance of link node v_i in the traffic network. $E(G)$ is used to evaluate the overall efficiency of the traffic network, with a larger $E(G)$ indicating a more efficient urban traffic network.

In the urban transportation network, the critical link nodes have a more significant impact on the overall traffic network efficiency. Therefore, after obtaining the importance of each road link, a certain percentage (such as 10%, 30%, 50%, or 70%) of critical road link nodes can be removed from the original network with an overall traffic network efficiency E to obtain the overall traffic network efficiency E' after removing the road link nodes. The traffic network efficiency ratio is calculated as follows:

$$\eta = \frac{E'}{E} \tag{10}$$

where E is the original traffic network efficiency, E' is the traffic network efficiency after removing a certain percentage of critical road link nodes. This indicator indicates the change in urban traffic network efficiency. When removing the same proportion of link nodes, a smaller value of η indicates a greater impact of the removed link nodes on the traffic network, thereby emphasizing the higher importance of the removed road links in the urban traffic system.

4.3. Baselines

In order to verify the effectiveness of the proposed method, different methods are selected and evaluated in comparison with the proposed method in this study on a uniform dataset:

- Sorted by traffic flow (SF)

This method sorts all link nodes by their traffic flow, as it is widely acknowledged that road links with higher traffic flow are more crucial in the traffic system. By comparing this method, we can evaluate the effectiveness of using solely traffic metrics for link importance evaluation. Additionally, we can investigate how the representation learning and clustering process impact the final results.

- Degree and clustering coefficients index (DCI)

This method based on degree and clustering coefficient was proposed by Yan et al. [26] for measuring the node importance evaluation of aviation networks. The method combines the direct influence of the nodes in the network and the closeness of the connection between the node neighbors. The formula of DCI of link node v_i in the traffic network is shown as follows:

$$DCI(i) = k_i \times \alpha^{-C(i)} (\alpha > 1) \tag{11}$$

where k_i is the degree of link node v_i, $C(i)$ is the clustering coefficient of link node v_i, α is an adjustable parameter, which is set to 3 in this paper. In this method, the larger DCI means the link node is more important in the traffic network.

- Deepwalk

Deepwalk, proposed by Perozzi et al. [21] in 2014, utilizes the concept of word2vec to generate node sequences by uniformly sampling nodes through random walk in the graph structure. The resulting node sequences are used as the corpus for word2vec, combined with the skip-gram model to learn the d-dimensional embedding vectors of the nodes. The obtained embedding vectors are then clustered using the K-means method, and their importance in each category is evaluated based on the importance evaluation indicator. The average importance evaluation indicator of each category reflects the importance of road links within that category, allowing for the identification of critical road link categories.

- Node2vec

Node2vec, proposed by Grover et al. [22] in 2016, modifies the node sampling method of Deepwalk by incorporating BFS and DFS. The resulting node sequences are then used as a corpus for the word2vec algorithm to learn d-dimensional embedding vectors of the nodes using the skip-gram model. The obtained embedding vectors are then clustered using the K-means method, and their importance in each category is evaluated based on the importance evaluation indicator. The average importance evaluation indicator of each category reflects the importance of road links within that category, allowing for the identification of critical road link categories.

5. Experimental Results

5.1. Parameter Sensitivity Analysis

The TraLink2vec model is composed of four key hyperparameters: p, q, α, and β. To investigate the impact of these hyperparameters on the clustering results, we analyze different time periods (morning and evening rush hours) and time intervals (T = 5 min, T = 15 min, T = 30 min). For simplicity, we fixed the dimension of embedding vectors ($d = 3$), random walk length ($l = 50$), the number of walks ($r = 10$), and window size ($w = 5$). Research data were collected on 12 January 2022.

Figure 4 shows the change of the CHI with the change of the return parameter p and in-out parameter q in static attributes. The figure shows that as the parameter q increases, the CHI decreases regardless of the time period, suggesting that smaller values of q result in better clustering results with higher CHI values. Moreover, the different colored curves in the figure correspond to different values of the parameter p. As the parameter p increases, the overall CHI curve shifts downward, indicating that CHI becomes smaller, and the clustering results worsen. This indicates that smaller values of p lead to better clustering results with higher CHI values.

Figure 5 displays the variation of CHI as the static and dynamic characteristic control parameters α and β change. These parameter variations reflect the model's attention to the traffic network structure and traffic information during the random walk. For T = 5 min and T = 30 min, the CHI curve generally increases as the static control parameter α increases, indicating that larger values of α lead to better clustering effects and higher CHI values. However, for T = 15 min, the curve shows fluctuations, indicating that parameter adjustments are necessary to achieve optimal results. The curves of different colors in the figure represent different parameters β. In general, the models obtain better clustering effects when $\beta < 1$, and no significant shift in the curve is observed when $\beta > 1$. The same dynamic control parameter β has different CHI values at different time intervals during different time periods, indicating that the effect of clustering changes with time periods and intervals. These findings demonstrate that the proposed model successfully captures the dynamic features in the urban traffic system during the learning of road link information. In practical applications, the optimal β value needs to be determined by adjusting the parameters.

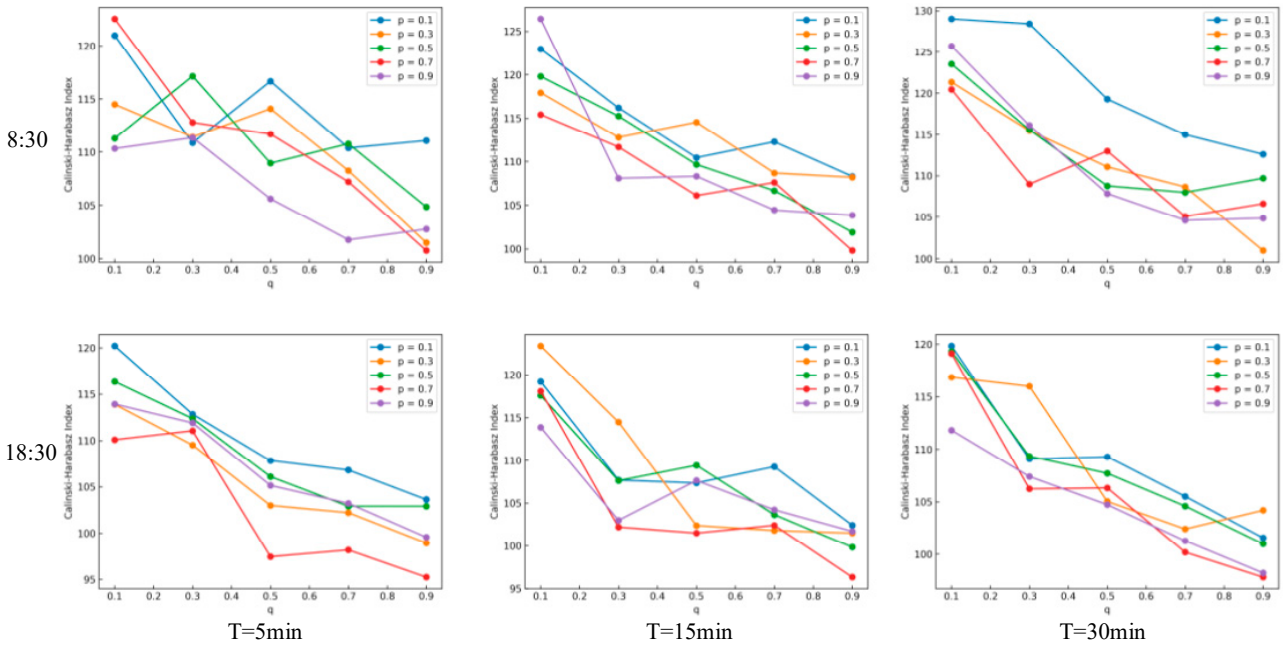

Figure 4. Parameter sensitivity of p and q.

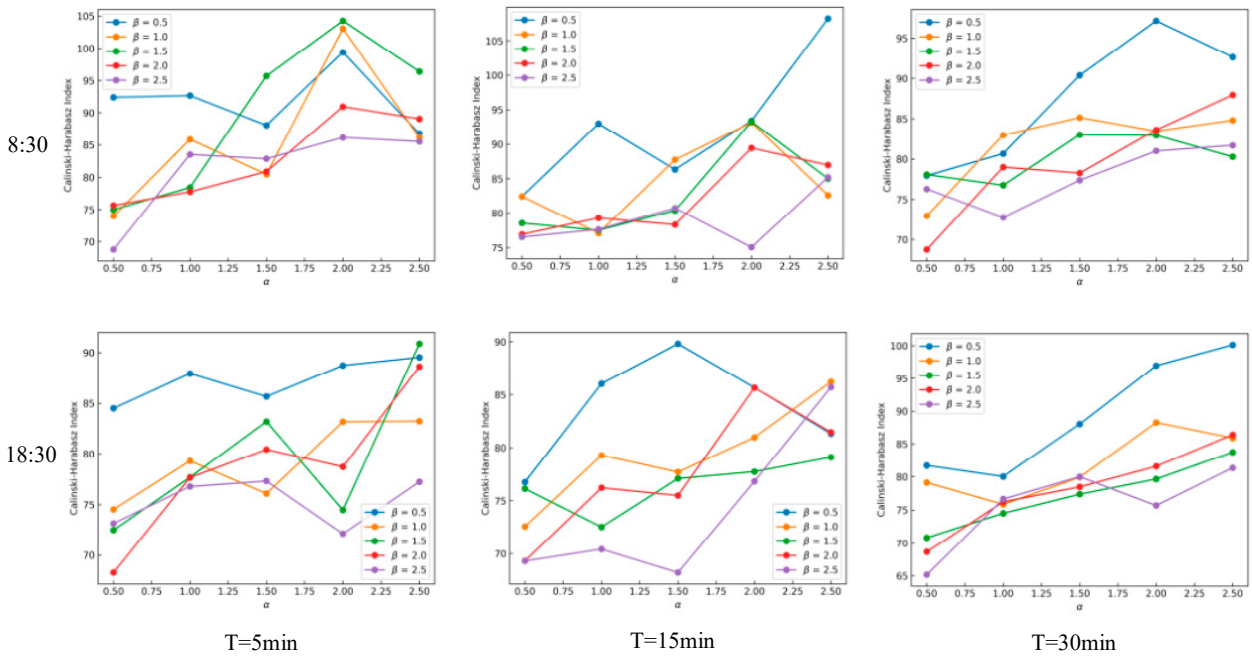

Figure 5. Parameter sensitivity of α and β.

5.2. Performance and Comparison

To evaluate the effectiveness of the proposed method, we compared four road link importance evaluation methods, including the sorted by traffic flow method (SF), degree and clustering coefficients index (DCI), Deepwalk, and Node2vec. A directed dual traffic network model was constructed using urban AVI data with different time periods and different time interval sizes, which is shown in Table 1. All data were collected on 12 January 2022.

Table 1. Illustration of different time periods and different time intervals.

Time	During Time		
	5 min	**15 min**	**30 min**
Morning Rush	8:30–8:05	8:30–8:45	8:30–9:00
Hollow	10:30–10:05	10:30–10:45	10:30–11:00
Evening Rush	18:30–18:05	18:30–18:45	18:30–19:00

This study selected the road links in the top 33% of importance as destructible road links for the importance evaluation using different methods. Moreover, for the destructible road links, the same proportion (10%, 30%, 50%, 70%) of link nodes were removed using random selection. The performance of the method was evaluated by comparing the magnitude of the traffic network efficiency ratio η of the urban road network before and after the road link removal. Where the smaller road network efficiency ratio η means that its method performs better. Each experiment was repeated 10 times to ensure the validity of the results in this paper.

For the selection of the set of destructible road links, the study obtained them in two ways:

(1) For the method based on importance ranking (SF, DCI), the study used the nodes ranked in the top 33% as the set of destructible road links.
(2) For the graph representation learning model (TraLink2vec, Deepwalk, Node2vec), the selection was based on the clustering results. If the number of link nodes in the category of critical links was greater than 33% of the total number of link nodes, the link nodes were randomly removed from it. If the number of link nodes was less than 33% of the total number of link nodes, the remaining link nodes were randomly selected from the category of normal road links after all the link nodes in the category of critical road links were selected.

For TraLink2vec, we set parameter α to 0.5, parameter β to 1, parameter p to 1, parameter q to 0.25, and maintained the other parameter settings as previously defined. For Deepwalk and Node2vec, the same parameter settings were used as in TraLink2vec.

Tables 2–4 present the performance of various methods for different time intervals during different time periods. The bold numbers indicate the optimal performance for each condition. The results show that the proposed methods in this study demonstrate superior performance in most cases, indicating that the integration of dynamic and static features of road links can effectively identify critical road links.

Table 2. Performance of road link node importance evaluation in different methods (T = 5 min).

Time	Proportion of Links Removed	The Traffic Network Efficiency Ratio				
		DCI	**SF**	**Deepwalk**	**Node2vec**	**TraLink2vec**
Morning Rush	10%	0.883	0.912	0.836	0.861	**0.785**
	30%	0.727	0.74	0.635	0.611	**0.604**
	50%	0.392	0.568	0.452	0.477	**0.406**
	70%	0.288	0.467	**0.266**	0.303	0.268
Hollow	10%	0.939	0.915	0.868	0.872	**0.859**
	30%	0.677	0.76	0.691	**0.545**	0.673
	50%	0.614	0.588	0.479	0.471	**0.466**
	70%	0.44	0.454	0.444	0.412	**0.271**
Evening Rush	10%	0.857	0.925	0.844	0.844	**0.837**
	30%	0.786	0.648	0.66	0.684	**0.639**
	50%	0.537	0.585	**0.465**	0.477	0.523
	70%	0.417	0.396	0.402	0.395	**0.318**

Table 3. Performance of road link node importance evaluation in different methods (T = 15 min).

Time	Proportion of Links Removed	The Traffic Network Efficiency Ratio				
		DCI	SF	Deepwalk	Node2vec	TraLink2vec
Morning Rush	10%	0.917	0.946	0.915	0.82	**0.805**
	30%	0.683	0.714	0.607	0.697	**0.442**
	50%	0.558	0.578	0.439	**0.379**	0.384
	70%	0.382	0.43	0.312	0.326	**0.292**
Hollow	10%	0.917	0.918	**0.846**	0.917	0.865
	30%	0.65	0.741	0.653	0.7	**0.611**
	50%	0.583	0.571	0.562	0.455	**0.477**
	70%	0.45	0.501	0.443	0.436	**0.376**
Evening Rush	10%	0.877	0.952	0.9	**0.826**	0.848
	30%	0.676	0.766	0.709	0.643	**0.573**
	50%	0.49	0.611	0.542	0.473	**0.431**
	70%	0.378	0.462	0.421	0.392	**0.332**

Table 4. Performance of road link node importance evaluation in different methods (T = 30 min).

Time	Proportion of Links Removed	The Traffic Network Efficiency Ratio				
		DCI	SF	Deepwalk	Node2vec	TraLink2vec
Morning Rush	10%	0.876	0.923	0.82	0.88	**0.801**
	30%	0.739	0.778	0.587	0.593	**0.483**
	50%	0.526	0.573	0.432	0.455	**0.366**
	70%	0.319	0.391	0.331	0.297	**0.274**
Hollow	10%	0.946	0.931	**0.854**	0.906	0.857
	30%	0.773	0.798	0.674	0.662	**0.599**
	50%	0.553	0.596	0.527	0.509	**0.505**
	70%	0.403	0.472	0.486	0.404	**0.261**
Evening Rush	10%	0.918	0.936	0.898	0.907	**0.845**
	30%	0.744	0.77	0.682	0.647	**0.592**
	50%	0.517	0.63	0.497	0.488	**0.465**
	70%	0.377	0.437	0.43	0.408	**0.284**

Compared to SF and DCI, the TraLink2vec model outperforms both methods in all cases, demonstrating that relying solely on traffic travel or network topology indicators cannot effectively identify critical road links in urban transportation systems. Compared to Deepwalk and Node2vec, although TraLink2vec did not achieve the best performance during short intervals (T = 5 min) and hollow periods, the difference between its performance and the best method can be negligible. This suggests that TraLink2vec is capable of capturing the latent information of urban structures even during periods with no apparent traffic travel characteristics and is also effective in identifying critical road links.

As the time interval increases, TraLink2vec demonstrates superior performance compared to other methods in identifying critical road links. Particularly in the evaluation with a time interval of 30 min, TraLink2vec significantly outperforms other baseline methods. This suggests that as the time interval increases, TraLink2vec, utilizing both the network structure and dynamic travel, can capture the hidden features in the traffic system and accurately identify critical road links in the city through the clustering method. Furthermore, as the proportion of removed link nodes increases, the performance gap between TraLink2vec and other methods widens. When the proportion of removed link nodes reaches 70%, the traffic network efficiency ratio of TraLink2vec is approximately between 0.25 and 0.3, indicating that removing less than 25% of link nodes will cause the network to collapse. On the other hand, when other methods with the same proportion of link nodes are removed, the traffic network efficiency is between 0.3 and 0.5. These results

demonstrate the advantages of the proposed method in accurately identifying critical road links in the traffic system.

6. Conclusions

The evaluation of urban critical links is of great significance for the control of urban traffic systems. In this paper, we propose a novel approach that combines both static and dynamic attributes of urban road networks to evaluate critical road links. The directed dual traffic network model is introduced to reflect the crucial role of road links in the traffic system. By constructing the directed dual traffic network at different time periods, the circulation and time-varying characteristics of the traffic system are fully considered. The TraLink2vec model is proposed to accurately capture the dynamic and static information of road links in the urban traffic network, which can effectively identify critical road links in urban road networks combined with machine learning analysis. The experimental results show that our proposed method outperforms baseline methods under real-world data, demonstrating its effectiveness and superiority.

Individual travel information is an important feature in traffic systems, which contains great spatio-temporal information. In the future, the combination of individual travel information and traffic travel information in urban areas has the potential to enable joint control of road link node information mining, leading to a comprehensive understanding of urban traffic. Additionally, the TraLink2vec model holds promise for application beyond critical road link evaluation, as it can be extended to explore the prediction of traffic travel information, including traffic status and travel trajectory, through combination with deep learning models.

Author Contributions: Conceptualization, Y.L. and M.H.; methodology, Y.L. and M.H.; software, Y.L.; validation, Y.L. and M.H.; formal analysis, Y.L. and M.H.; investigation, Y.L. and M.H.; resources, M.H.; data curation, Y.L.; writing—original draft preparation, Y.L.; writing—review and editing, Y.L. and M.H.; visualization, Y.L.; supervision, M.H.; project administration, Y.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Key Research and Development Program of China, grant number 2020YFB1600400.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: The authors would like to thank the support of the National Key Research and Development Program of China (2020YFB1600400). The authors also wish to thank anonymous referees for their valuable comments.

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

References

1. Daqing, L.; Yinan, J.; Rui, K.; Havlin, S. Spatial Correlation Analysis of Cascading Failures: Congestions and Blackouts. *Sci. Rep.* **2014**, *4*, 5381. [CrossRef] [PubMed]
2. Çolak, S.; Lima, A.; González, M.C. Understanding Congested Travel in Urban Areas. *Nat. Commun.* **2016**, *7*, 10793. [CrossRef] [PubMed]
3. Louail, T.; Lenormand, M.; Cantu Ros, O.G.; Picornell, M.; Herranz, R.; Frias-Martinez, E.; Ramasco, J.J.; Barthelemy, M. From Mobile Phone Data to the Spatial Structure of Cities. *Sci. Rep.* **2014**, *4*, 5276. [CrossRef] [PubMed]
4. Goodchild, M.F.; Janelle, D.G. The City around the Clock: Space-Time Patterns of Urban Ecological Structure. *Environ. Plan. A* **1984**, *16*, 807–820. [CrossRef]
5. Zhao, S.; Zhao, P.; Cui, Y. A Network Centrality Measure Framework for Analyzing Urban Traffic Flow: A Case Study of Wuhan, China. *Phys. A Stat. Mech. Appl.* **2017**, *478*, 143–157. [CrossRef]
6. Tian, Z.; Jia, L.; Dong, H.; Su, F.; Zhang, Z. Analysis of Urban Road Traffic Network Based on Complex Network. *Procedia Eng.* **2016**, *137*, 537–546. [CrossRef]

7. Saberi, M.; Mahmassani, H.S.; Brockmann, D.; Hosseini, A. A Complex Network Perspective for Characterizing Urban Travel Demand Patterns: Graph Theoretical Analysis of Large-Scale Origin–Destination Demand Networks. *Transportation* **2017**, *44*, 1383–1402. [CrossRef]
8. Girvan, M.; Newman, M.E. Community Structure in Social and Biological Networks. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 7821–7826. [CrossRef]
9. Wang, J.; Rong, L.; Guo, T. A New Measure Method of Network Node Importance Based on Local Characteristics. *J. Dalian Univ. Technol.* **2010**, *50*, 822–826.
10. Tu, Y.; Yang, C.; Chen, X. Analysis of Road Network Topology Vulnerability and Critical Links. *J. Tongji Univ. Nat. Sci.* **2010**, *38*, 364–367+379.
11. Wang, L.; Huang, M.; Gao, L. Methods of Importance Evaluation of Traffic Network Node Based on Clustering Algorithms. *J. Transp. Inf. Saf.* **2020**, *38*, 80–88.
12. Yi-Run, R.; Song-Yang, L.; Jun-De, W.; Liang, B. Ranking Node Importance in Large-Scale Complex Network: From a Perspective of Local Abnormal Links. In Proceedings of the 2017 3rd International Conference on Big Data Computing and Communications (BIGCOM), Chengdu, China, 10–11 August 2017.
13. Su, F.; Dong, H.; Jia, L.; Sun, X. Identification of Critical Section in Urban Traffic Road Network Based on Space-time Correlation. *J. Transp. Syst. Eng. Inf. Technol.* **2017**, *17*, 213–221.
14. Chen, B.Y.; Lam, W.H.K.; Sumalee, A.; Li, Q.; Li, Z.-C. Vulnerability Analysis for Large-Scale and Congested Road Networks with Demand Uncertainty. *Transp. Res. Part A Policy Pract.* **2012**, *46*, 501–516. [CrossRef]
15. Scott, D.M.; Novak, D.C.; Aultman-Hall, L.; Guo, F. Network Robustness Index: A New Method for Identifying Critical Links and Evaluating the Performance of Transportation Networks. *J. Transp. Geogr.* **2006**, *14*, 215–227. [CrossRef]
16. Jenelius, E.; Mattsson, L.-G. Road Network Vulnerability Analysis of Area-Covering Disruptions: A Grid-Based Approach with Case Study. *Transp. Res. Part A Policy Pract.* **2012**, *46*, 746–760. [CrossRef]
17. Sun, J.; Xiang, Q.J. Identification of Critical Links Based on the Failure Consequence Evaluation. In Proceedings of the 2018 3rd IEEE International Conference on Intelligent Transportation Engineering (ICITE), Singapore, 3–5 September 2018.
18. Zhang, J.; Jia, L.; Niu, S.; Li, H. Identification and analysis of road network key segments set based on K-shortest path. *J. Chang. Univ. Nat. Sci. Ed.* **2015**, *35*, 122–129.
19. Porta, S.; Crucitti, P.; Latora, V. The Network Analysis of Urban Streets: A Primal Approach. *Environ. Plan. B Plan. Des.* **2006**, *33*, 705–725. [CrossRef]
20. Porta, S.; Crucitti, P.; Latora, V. The Network Analysis of Urban Streets: A Dual Approach. *Phys. A Stat. Mech. Appl.* **2006**, *369*, 853–866. [CrossRef]
21. Perozzi, B.; Al-Rfou, R.; Skiena, S. DeepWalk: Online Learning of Social Representations. In Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, New York, NY, USA, 24–27 August 2014.
22. Grover, A.; Leskovec, J. Node2vec: Scalable Feature Learning for Networks. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, 13–17 August 2016.
23. Hartigan, J.A.; Wong, M.A. Algorithm AS 136: A K-Means Clustering Algorithm. *Appl. Stat.* **1979**, *28*, 100–108. [CrossRef]
24. Caliński, T.; Harabasz, J. A Dendrite Method for Cluster Analysis. *Commun. Stat.* **1974**, *3*, 1–27.
25. Mo, H.; Deng, Y. Identifying Node Importance Based on Evidence Theory in Complex Networks. *Phys. A Stat. Mech. Appl.* **2019**, *529*, 121538. [CrossRef]
26. Yan, L.; Chen, Z.; Zhang, Q. Analysis of Key Nodes in China's Aviation Network Based on the Degree Centrality Indicator and Clustering Coefficient. *CAAI Trans. Intell. Syst.* **2016**, *11*, 586–593.