

Bike allocation strategies in a competitive dockless bike sharing market

Abstract: This paper investigates bike allocation strategy in a competitive dockless bike sharing market from both market leader and market follower's perspective. Market leader aims to grasp the market using the shortest time, while market follower wants to get the market share from market leader. The initial bike allocation position, coverage area and allocation method are investigated based on community structure method. An approximate optimal allocation model is proposed to maximise the recourses. Numerical examples are tested in a simplified Sioux Falls network and a real bike sharing network in Singapore to illustrate the effectiveness and practicability. Results show that the proposed method could identify the reasonable and efficient allocation position quickly. Market follower should apply the same allocation strategy as market leader, while use lower ticket, more bikes, and large batches allocation to strive for market share.

Keywords: dockless bike sharing; allocation strategy; complex network; community structure; competitive market

1. Introduction

Bike sharing is a sustainable transport service to provide the travellers a public bike for a short time. The development of bike sharing system has gone through from first generation free bikes to the current fourth generation dockless bike sharing (Meng et al., 2016; Zhang et al, 2017). Many cities currently have both station-based bike sharing system and dockless bike sharing system in the market. Station-based bike sharing systems are usually operated or regulated by government agency, while dockless bike sharing systems are all operated by privacy firms. Ofo, the biggest dockless bikes haring operator, has deployed over 10 million bikes in 250 cities and 20 countries.

Table 1 lists the typical studies on bike sharing systems in recent years. It could be found that compared to the mature research on the station-based bike sharing issues, the study on dockless bike sharing is mainly about demand analysis due to the recent boom. As it is a new service in the market, the demand is hardly perfectly predicted. Moreover, oversupply leads to a waste of resources in terms of money and road space. Considering the limited budget and unpredictable demand, it is worth to investigate how to allocate the bikes efficiently to obtain maximum benefits? Where and how many bikes should the operator be planned? How large area could these bikes be covered? Vogel et al. (2014) presented a hybrid metaheuristic method to relocate the bikes in a station-based bike sharing system. Method performed better than mathematical algorithms using the Vienna's data. Similar algorithms have been proposed by other scholars to achieve the balance status (García-Palomares et al., 2012; Raviv and Kolka, 2013). Meng et al. (2015) proposed a Geographic Information System (GIS)-based cellular automata (CA) model for planning a new bike sharing network (link, node, docking station) within an existing transport network. A model prototype was developed in a common desktop GIS and applied to a rapidly developing area in Singapore. The maximal coverage issue has been also studied by Frade and Ribeiro (2015) considering budget constraints, Liu et al. (2015) by considering multiple influential factors, and Shu et al. (2019) by considering the optimal distance between the building entrance/exit and the station. Datner et al. (2017) developed a robust guided local search algorithm to set the target inventory level for bike relocation. Kabak et al. (2018) proposed a method to locate future bike stations in an existing bike sharing system using multi-criteria decision-making method on a GIS platform. Chen et al. (2018) investigated how to determine the number of bikes to be deployed at each station during bike repositioning in bike-

sharing systems. The objective function is to maximise the time interval between repositioning events. The above studies all targeted on station-based bike sharing system, while these investigations have not been discussed in dockless bike sharing system. Such research findings can be helpful in provide guidance in planning dockless bike sharing schemes.

Table 1 Typical studies on bike sharing system

Authors	Year	System type	Research problem	Practical application
Ho and Szeto	2014	Station-based	Bike reposition	No
Forma et al.	2015	Station-based	Bike reposition	No
Zhang et al.	2015	Station-based	Demand analysis	Yes, five cities in China
Bordagaray et al.	2016	Station-based	Demand analysis	Yes, Santander
Szeto et al	2016	Station-based	Bike reposition	Yes, Vienna
Fishman	2016	Both	Review	No
Faghieh-Imani et al.	2017	Station-based	Bike reposition	Yes, Barcelona
Pal and Zhang	2017	Dockless	Bike reposition	No
Hsu et al.	2018	Station-based	Service improvement	Yes, Taipei
Shen et al.	2018	Dockless	Demand analysis	Yes, Singapore
Zhou and Ni	2018	Dockless	Demand analysis	Yes, Shanghai
Pan et al.	2018	Dockless	Bike reposition	Yes, Shanghai
Zhang et al.	2019	Dockless	Fence planning	Yes, Shanghai

This study aims to propose the allocation strategy for both market leader and market follower in a competitive market using community structure method in complex network. Community structure is one common characteristic of complex network. A group of vertices in a network is easily to cluster together with high concentrations of edges and generate a structure in a hierarchical way. This phenomenon in real network is defined as community structure or clustering (Girvan and Newman, 2002). Community, also known as cluster or module, is a group of vertices which have highly similarities in the network. The community structure makes the generation and evolution of networked system much quicker and more stable than if the system is unstructured (Zhou et al, 2012). One of the most classical algorithms is Fast Newman (F-N) algorithm, which is a bottom-up agglomerative algorithm (Newman, 2004). This method starts from a set of isolated vertices and then iteratively calculate the modularity by adding the vertices in the group. The advantage of using community structure theory in bike allocation issue is because: (1) Dockless bike sharing system exists in urban transport system, which also has the feature of community structure. Bike positions could be regarded as vertices and bike routing could be regarded as edges. (2) Due to the characteristics of short-distance usage, bikes are easily distributed within a fixed area near the initial allocation position. Bikes may not be distributed to another area with long distance travel shortly. Therefore, the community structure method is suitable to be used in bike allocations problem. It is also efficient to use the complex network concept in dealing with a large-scale bike sharing system.

This study contributes to the existing literature in terms of: (1) this study systematically developed a bike allocation framework in dockless bike sharing system in a competitive market, where both market leader and follower have been considered. The appropriate operation will

help the market development on a sustainable track. (2) this is the first attempt to adopt community structure method into bikes sharing operation; (3) this study proposed an approximate optimal bike allocation model based on the bike allocation framework.

2. Basic definition

As shown in Figure 1, vertices within the community are densely connected; while between communities are less connected. In dockless bike sharing system, when the bikes are placed at one vertex, due to the stochastic characteristic of demand, these public bikes shall be distributed to other vertices. Different local features (e.g. demand level, land use characteristics, route direction and geography constraint) will lead to different coupling degrees among vertices. Thus, community structure will be formed by clustering the vertices with similar coupling degree. Vertices in the same community are densely connected with high traffic volume, while vertices in different communities are sparsely connected with low or even zero traffic volume.

There are various types of methods to detect the community structure. Among them, modularity-based method has widely adopted in large network application due to its fast and efficient computation. The most typical modularity-based method is greedy method proposed by Newman (2004), which is an agglomerative hierarchical clustering method. Applying it into a dockless bike sharing system, the network could be denoted by a complex network $G = (V, E, W)$, where V, E, W represent the set of vertex, edge and weight, respectively. There are N number of vertices and E number of edge in G . Several key concepts are defined as follows in a dockless bike sharing network.

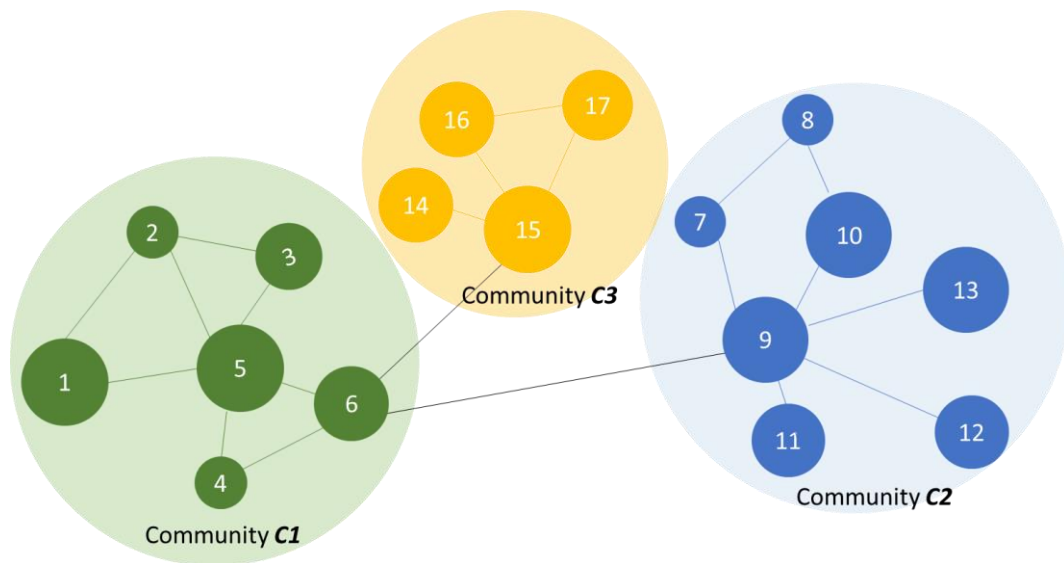


Figure 1 A sketch of a small network with three communities

Weight of edge: w_{uv} is the weight of direct edge from vertex u to v . The higher value of w_{uv} reflects the stronger correlation between vertices u and v . Conversely, the lower value of w_{uv} indicates the weaker correlation between vertices u and v .

Weight of vertex: ζ_u is the weight of vertex u , which is the sum of the weight of the connected edges, that is $\zeta_u = \sum_v w_{uv}$. Higher weight of vertex means that this vertex has a significant influence on the network. In an unweighted network, $w_{uv}=1$, the weight of vertex is equal to the degree of vertex.

Degree of vertex: k_u is the degree of vertex u . Each edge in the network connects two vertices, which are neighbours of each other. All the neighbours of one vertex is called the vertex's neighbourhood. The number of neighbours in the neighbourhood is the degree of the vertex. Larger degree of vertex means that the vertex has more neighbours and therefore it has a higher importance in the network.

Modularity: M is a quality function to test whether a clustering is meaningful in community structure theory. Higher modularity of the network means that there is a clear community structure in the network. The concept of modularity is proposed based on coordination theory which is firstly proposed by Newman (2004). The philosophy is to compare the network after partition with a null model, which has the same configuration as the original network but contains no community structure because of a random placement of its edges (Chang et al., 2012).

In an unweighted network, set A_{uv} is an element in the adjacent matrix in the network. If vertex u is linked with vertex v , $A_{uv} = 1$; otherwise, $A_{uv} = 0$. Assume vertices u and v are belong to the communities of c_u and c_v . The ratio of number of edges in the community to the number of edges in the network is calculated by:

$$\frac{\sum_{uv} A_{uv} \delta(c_u, c_v)}{\sum_{uv} A_{uv}} = \frac{1}{2m} \sum_{uv} A_{uv} \delta(c_u, c_v) \quad (1)$$

where $\delta(c_u, c_v)$ is the relationship function of c_u and c_v . If vertices u and v are belong to the same community, that is $c_u = c_v$, then $\delta(c_u, c_v) = 1$. Otherwise, $\delta(c_u, c_v) = 0$. m is the total number of edges in the network. Modularity M is defined as the fraction of edges within community minus the expected fraction of such edges in a randomised null model of the network (Yan and Gerstein, 2017). Mathematically, the modularity is equal to

$$M = \frac{1}{2m} \sum_{uv} \left[A_{uv} - \frac{k_u k_v}{2m} \right] \delta(c_u, c_v) \quad (2)$$

where k_u is the degree of vertex u .

$$k_u = \sum_v A_{uv} \quad (3)$$

Let e_{ij} be the ratio of the total number of edges in communities i and j to the total number of edges in the network, a_i be the ratio of number of edges connected with all the veticles in community i to the total number of edges in the network. That are,

$$e_{ij} = \frac{1}{2m} \sum_{uv} A_{uv} \delta(c_u, i) \delta(c_v, j) \quad (4)$$

$$a_i = \frac{1}{2m} \sum_u k_u \delta(c_u, i) \quad (5)$$

To simplify the calculation of the modularity M , assume the network has been divided into n ($i=1,2,\dots,n$) number of community. Eq.(2) could be express as:

$$\begin{aligned}
M &= \frac{1}{2m} \sum_{uv} \left[A_{uv} - \frac{k_u k_v}{2m} \right] \delta(c_u, c_v) \\
&= \frac{1}{2m} \sum_{uv} \left[A_{uv} - \frac{k_u k_v}{2m} \right] \sum_i \delta(c_u, i) \delta(c_v, i) \\
&= \sum_{i=1}^n \left[\frac{1}{2m} \sum_{uv} A_{uv} \delta(c_u, i) \delta(c_v, i) - \frac{1}{2m} \sum_u k_u \delta(c_u, i) \frac{1}{2m} \sum_v k_v \delta(c_v, i) \right] \\
&= \sum_{i=1}^n (e_{ii} - a_i^2)
\end{aligned} \tag{6}$$

3. Problem statement and algorithm

Some assumptions are made as follows:

- A1: Operator should minimise the number of bikes to optimal the resource allocation.
- A2: Operator could not predict precisely the demand and will allocate the bikes in bathes.
- A3: Due to the production time of bikes, operator only have small number of new bikes available for each allocation.
- A4: Operator will allocate the bikes repeatedly when the market is under supply.
- A5: Bikes will be distributed by riders via trips. The repositioning issue is not considered in this study.
- A6: There will be two operators in the market: market leader and market follower. The market leader has potential competitor – market follower, which shall enter into the market soon. Two scenarios are considered in this study as follows: Scenario A: from the market leader’s perspective, how to allocate the bikes to grasp the market? Scenario B: from the market follower’s perspective, how to enter the market to compete against the market leader? The time of competitive market entry for the three scenarios is illustrated in Figure 2.

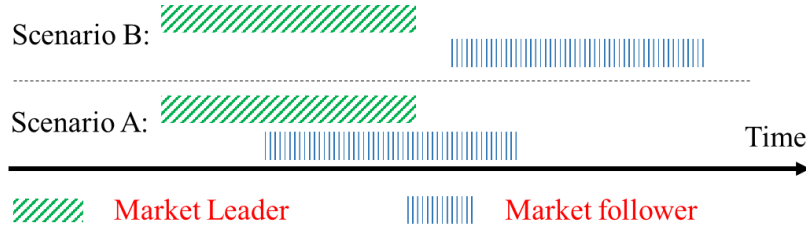


Figure 2 The time of competitive market entry for two scenarios

Since dockless bike sharing is an emerging market, the market leader may not precisely predict the rider demand. Also considering the production time and the operation budget, the most efficient way to allocation bikes is allocate in small batches to test the demand. Generally, an allocation strategy framework could be designed by an iterative process:

Step 1: initialise the system and set the basic parameters. Once the planning area is confirmed, certain number of TAZs need to be classified. The total demand for the study area and each TAZ is predicted according to the land use pattern.

Step 2: detect the community structure in the network and find the key vertex in each community. The features of the key vertex are defined by the operator according to its market objective.

Step 3: allocate certain number of bikes to the key vertex in each community.

Step 4: check whether the number of bikes could meet the demand requirement. If yes, the allocation is finished. Otherwise, update demand in each TAZ and go to step 2 to allocate new bikes.

Since each bike is equipped with GPS, operator could easily locate the bike and analyse the usage. The updated demand is equal to the difference between the total demand of each TAZ minus the balance bikes. Apparently, when a vertex always has available bikes during the whole day, it reflects that the bikes at this vertex is oversupplied.

3.1 Scenario A

Considering the potential competitor, as the market leader, the primary goal is to dominate the market before the competitor enters into the market. The objective of the market leader is to attract more riders in the shortest time. In the circumstances, the market leader is best to allocate the bikes at the positions which could distribute the bikes with the fastest speed. The bike distribution speed relays on the demand level of the vertex, which is the weight of the vertex. From a network (physical) perspective, the public bikes will be distributed to the whole area in the community with the fastest speed only when the placing position is the vertex with the highest weight in the network. As shown in Figure 1, if the network is unweighted, the allocation points are vertex 5, 9 and 15 in community c1, c2 and c3 respectively. This phenomenon has been discussed widely in the social network, called leader-follower relationship. The leader in the community structure is the vertex with the highest weight and plays the key role in that community (Preece and Shneiderman, 2009; Wu et al., 2013). In this scenario, the highest weight vertex is the allocation point, and relevant community is its coverage area.

Applying this objective in the allocation strategy design framework, a modified community structure detection algorithm is proposed. It is a NP hard problem to find the optimal solution to a network partitioning issue in a large-scale transport network. Due to the high complexity in computation, many researches have been published to ease the procedure and improve the accuracy of solution. The main steps of F-N algorithm are listed as follows:

Step 1: Network initialisation. Each vertex represents one community;

Step 2: Community combination. Two communities are joined together only if the join could result in the greatest increase in modularity;

Step 3: Repeat Step 2 until the whole network is jointed as one community;

Step 4: Optimal result selection. The optimal result of community structure is the best cut with the largest modularity in the process.

However, the F-N algorithm is proposed for unweighted network. To adopted it in the bike sharing network which has weightage on each edge, some adjustments were made. First, e_{ij} is redefined as:

$$e_{ij} = \frac{\sum_{uv} w_{uv} A_{uv} \delta(c_u, i) \delta(c_v, j)}{\sum_{uv} w_{uv}} \quad (7)$$

where w_{uv} is the edge weight. Applying the Gravity model, w_{uv} is proportional to the passenger demand between vertices u and v , and inversely proportional to the square of the distance

between vertices u and v . Meanwhile, cycling is not suitable for over short trips (e.g. distance less than 0.1km) or over long trips (e.g. distance more than 3km). the edge weight is limited as:

$$w_{uv} = \begin{cases} (p_u p_v) / (d_{uv}^2), & \forall d_{uv} \in [0.1, 3] \\ 0, & \text{otherwise} \end{cases} \quad (8)$$

where p_u and p_v are the passenger demand of vertex u and v , d_{uv} is the shortest distance between vertex u and v . In bike sharing network, an edge with higher weight means there is larger bike sharing flow on this edge.

Considering that weight of vertex and degree of vertex are both the indicators from vertex point of view, which is only one-side of the network evaluation, an improvement is added with taken the weight of edge, that is:

$$a_i = \gamma \sum_j e_{ij} + (1 - \gamma) \cdot k_i / (2m) \quad (9)$$

The detailed procedures are listed as follows:

Step 1: divide the whole study area into n number of TAZs;

Step 2: number all TAZs and convert the study area into an undirected network composed of vertices and edges;

Step 3: calculate the correlation weight w_{uv} of two adjacent TAZs u and v by using the method proposed in Eq.(8) as the weighted value of each edge in the network, and then construct the corresponding adjacency matrix;

Step 4: initialise the network with a state in which each vertex is the sole unit of one of n community;

Step 5: join communities together in pairs sequentially and compute the variation of modularity. ΔM given by the merging of any two communities of the running partition as follows:

$$\Delta M = e_{ij} + e_{ji} - 2a_i a_j = 2(e_{ij} - a_i a_j) \quad (10)$$

Step 6: choose the join that results in the greatest increase or smallest decrease in M on the basis of greedy algorithm, and then compute the current modularity;

Step 7: update the elements e_{ij} ;

Step 8: execute Steps 5–7 repeatedly until all communities merge into one community.

This algorithm would carry out $n-1$ merging operations until completion, and hence, the entire computational complexity is $O((m+n)n)$, or $O(n^2)$ on a sparse graph (Newman, 2004). This algorithm is simple to realise the optimal decision of the position.

3.2 Scenario B

This scenario considers the market competition. The objective of the market follower is to take the market share from the market leader. The allocation strategy for the market follower is that first use the same allocation positions with the market leader, and then introduce various incentive methods to attract riders, e.g. price competition, service improvement. This scenario also applies for the market leader who would like to take back lost market. The process of seizing this market between market leader and market following is similar to the process of virus spread and immune in complex network. In the theory of immune control, for the

unbalanced network, the most effective immune vertex is the vertex with the largest weight. Thus, the allocation strategy should be in line with the strategy that has been obtained in Scenario A. When market follower uses the same allocation strategy, it is hard to grab market share from the market leader due to the brand loyalty and inertia.

To compete with the market leader, the market follower must provide other benefits to attract riders. Since bike sharing services are with similar features in term of bike quality, usage method and service price. After operation for a while, as the demand could be reflected from the current usage, the most efficient feature that could distinguish the characteristic of operator is the convenience of bikes. That is whether the operator could provide enough bikes for riders at any time. Therefore, the market follower should allocate the bikes in large batches instead of small batch as market leader.

A general case is considered in this scenario, where there are $n-1$ number of market leaders and only one market follower. Considering that there will be N_u total demand in future, φ proportion of them are exiting users of bike sharing and $(1-\varphi)$ of them are potential riders. Assume the potential riders will select the different operators following a logit type model, the probability of choosing operator n is

$$P_n = \exp(u_n) / \sum_{i=1}^n \exp(u_i) \quad (11)$$

All the bikes from different operators are with similar features. The utility function u_i is related with monetary cost f_i and the accessibility. The accessibility is relevant to the number of bikes in the network. The more bikes, the easier access for the riders. Therefore, the utility function could be expressed by

$$u_i = -\alpha f_i + \beta Nv_i / \text{Avg}(Nv) \quad (12)$$

where Nv_i is the total number of public bikes of the new entrant, $\text{Avg}(Nv)$ is the average public bikes of other operators, α, β are the converting factors.

The potential users that could be attracted by the new entrant are calculated by

$$N_{np} = (1-\varphi) N_u \cdot \exp(u_n) / \sum_{i=1}^n \exp(u_i) \quad (13)$$

For the exiting users, as they already paid certain number of deposit to the established firm, they are less sensitive to the usage monetary cost of new entrant. Assume the existing users could only consider using the bikes from the new entrant only if the accessibility of the bikes is higher. the number of switching from other operators to the new entrant could be expressed by

$$N_i = \left(1 - \exp\left(\mathcal{G}\left(1 - \frac{Nv_i}{\text{Avg}(Nv)}\right)\right)\right) \varphi N_u \quad (14)$$

Therefore, the total number of users for new entrant is

$$P = \left(1 - \exp\left(\mathcal{G}\left(1 - \frac{Nv_i}{\text{Avg}(Nv)}\right)\right)\right) \varphi + \frac{\exp\left(-\alpha f_n + \beta \frac{Nv_i}{\text{Avg}(Nv)}\right)}{\sum_{i=1}^n \exp\left(-\alpha f_i + \beta \frac{Nv_i}{\text{Avg}(Nv)}\right)} (1-\varphi) \quad (15)$$

If the target market share of new entrant is \bar{P} , the new entrant should design the price and provide the number of bikes following

$$P = \left(1 - \exp \left(\mathcal{G} \left(1 - \frac{Nv_i}{\text{Avg}(Nv)} \right) \right) \right) \varphi + \frac{\exp \left(-\alpha f_n + \beta \frac{Nv_i}{\text{Avg}(Nv)} \right)}{\sum_{i=1}^n \exp \left(-\alpha f_i + \beta \frac{Nv_i}{\text{Avg}(Nv)} \right)} (1 - \varphi) \geq \bar{P} \quad (16)$$

Because the bikes are only allocated at several vertices, the Eq.(16) is the basic requirement to achieve the goal for these certain vertices. To explain whether the success of certain vertices could guarantee the success of other vertices, epidemic threshold is used.

In the theory of virus spread, whether the virus could be largely spread depends on the diffusion probability. Let η is the probability of the vertex from susceptible to infected, ξ is the probability of the vertex from infected to susceptible, the valid diffusion rate could be expressed as

$$\lambda = \eta / \xi \quad (17)$$

When $\eta \leq \xi$, virus could not be spread out. Similarly, if the established firm has higher capability to occupy the market, the public bikes that are belong to the new entrant could not be distributed to other vertices after allocating. It means these new bikes will not be used by riders.

Assume the network is a scale-free network, let $\rho(t)$ is the density of success vertices at time t. when t approaches to infinity, the density is stable as ρ . Based on mean-field theory, the density equation of success vertices is

$$\partial \rho_k(t) / \partial t = -\rho_k(t) + \lambda k [1 - \rho_k(t)] \Theta(\rho_k(t)) \quad (18)$$

When $\rho(t) \rightarrow 1$ and set the right part in the equation is equal to zero, the steady state density is

$$\rho_k = \frac{\lambda k \Theta(\lambda)}{1 + \lambda k \Theta(\lambda)} \quad (19)$$

It means that the higher the modularity of the vertex, the higher probability to success. Considering the unbalanced transportation network, the probability of any edge direction to vertex s is $sP(s) / \langle k \rangle$, then it could obtain

$$\Theta(\lambda) = \frac{1}{\langle k \rangle} \sum_k k P(k) \rho_w \quad (20)$$

Combine Eqs.(16) and (17), when Θ is sufficiently small, ρ_w and $\Theta(\lambda)$ could be calculated approximately. When the diffusion rate is larger than an epidemic threshold $\lambda > \lambda_c$, Θ has a non-zero solution.

$$\Theta(\lambda) = \frac{1}{\langle k \rangle} \sum_k k P(k) \frac{\lambda k \Theta(\lambda)}{1 + \lambda k \Theta(\lambda)} \quad (21)$$

Where $\Theta = 0$ is the trivial solution. When $\Theta \neq 0$, it requires

$$\frac{d}{d\Theta} \left(\frac{1}{\langle k \rangle} \sum_k kP(k) \frac{\lambda k \Theta(\lambda)}{1 + \lambda k \Theta(\lambda)} \right) \Big|_{\Theta=0} \geq 1 \quad (22)$$

That is

$$\left(\sum_k \frac{kP(k) \lambda k}{\langle k \rangle} \right) = \frac{\langle k^2 \rangle}{\langle k \rangle} \lambda \geq 1 \quad (23)$$

Therefore, the critical value of the diffusion rate is

$$\lambda_c = \langle k \rangle / \langle k^2 \rangle \quad (24)$$

Because the basic requirement is $\lambda \geq \lambda_c$, the place vertices could distribute the bikes successfully. Based on Eq. (21), when the network is large enough $N \rightarrow \infty$ and $\langle k^2 \rangle \rightarrow \infty$, then $\lambda_c \rightarrow 0$, $\lambda \geq \lambda_c$ could be satisfied easily. It means that the operator can easily achieve the goal when the certain vertices are successful.

4. Extension

Section 3 gives the bike allocation strategies for each scenario in a step by step procedure, which may not be the overall optimal allocation plan. However, it could be extended to a bi-level approximate optimal model to guide the overall bike allocation. The approximate optimal solution is also important for operators as it can reduce the cost using shorter computing time. The detailed procedure is introduced as follows:

For a fixed area, population and demand are also fixed. Correspondingly, the total number of allocated bikes is fixed. The approximate optimal model is to determine the allocation times and the allocated bikes in each time with minimum operation cost. Therefore, the lower level mode is the strategy in section 3 while the upper model is to minimise the operation cost. The operation cost during bike allocation includes the management cost (e.g. manpower cost, vehicle routing cost, allocation station visit cost) and production cost (e.g. normal production cost and extra hour work for speeding up the production). Meanwhile, the fare revenue from different strategies are also different. Therefore, the cost function could be expressed as:

$$\text{Cos} = AC + PC - SR \quad (25)$$

where AC is the management cost, PC is the production cost, SR is the fare revenue. AC is relevant with the total number of allocated bikes and the total number of allocated station, which could be described as:

$$AC = \sum_{t=1}^T c^{al} (\mu_1 n_t^{stop} + \mu_2 n_t^{al,bike}) \quad (26)$$

where t is the allocation day, c^{al} is one unit of allocation cost, n_t^{stop} and $n_t^{al,bike}$ are the number of allocated station and bikes on that day, μ_1 and μ_2 are the qualified parameters.

The production cost is related to the number of required allocated bikes. As the production capacity is fixe, if the required allocated bikes are greater than the capacity, more resources are needed to increase the production. Then the production cost per each bike increased. Assuming that the total bike production is proportional to the unit cost of one bike, that is:

$$c_t^{bike} / c^{bike} = n_t^{p,bike} / n^{p,bike}, \forall n_t^{p,bike} \geq n^{p,bike} \quad (27)$$

where c^{bike} and c_t^{bike} are the normal production cost per bike and the specific production cost per bike on day t, respectively. $n^{p,bike}$ and $n_t^{p,bike}$ are the bike production number for normal day and day t. Therefore, the production cost could be expressed as:

$$PC = \sum_{t=1}^T c_t^{bike} n_t^{p,bike} \quad (28)$$

The fare revenue could be obtained based on the average usage time and the fare per each time:

$$SR = \sum_{t'=2}^T \sum_{t=2}^{t'} n_{t-1}^{al,bike} n^u p \quad (29)$$

where n^u is the average usage time per each bike, p is the fare per each time.

There are also several bike constraints as follows:

$$D \cdot n^{per} = N \cdot n^u \quad (30)$$

$$\sum_{t=1}^T n_t^{al,bike} = N \quad (31)$$

$$\sum_{t=1}^{T-1} n_t^{p,bike} = N \quad (32)$$

$$\sum_{t=1}^{t'} n_t^{p,bike} \geq \sum_{t=2}^{t'+1} n_t^{al,bike} \quad (33)$$

Eq.(30) represents that the total number of allocated bike can meet the total bike travel demand in the area, wherein D is the total bike travel persons, n^{per} is the average cycle frequency per person, N is the total number of allocated bikes, n^u is the average usage time per bike. Eqs.(31) and (32) mean that the total number of allocated and produced bikes can meet the required bike amount. Eq. (33) is the accumulated allocated bikes could not be greater than actual produced number of bikes. Other restraints also include minimum production limitation and non-negative limitation.

The upper level objective function in scenario A could be expressed as follows, wherein the market leader aims to grab the market as soon as possible:

$$Min \text{ Cos} = \left[\sum_{t=1}^T c_t^{bike} n_t^{p,bike} + \sum_{t=1}^T c^{al} (\mu_1 n_t^{stop} + \mu_2 n_t^{al,bike}) - \sum_{t'=2}^T \sum_{t=2}^{t'} n_{t-1}^{al,bike} n^u p \right] T \quad (34)$$

The upper level objective function in scenario B could be expressed as follows, wherein the market leader is to cover the maximum area:

$$Min \text{ Cos} = \left[\sum_{t=1}^T c_t^{bike} n_t^{p,bike} + \sum_{t=1}^T c^{al} (\mu_1 n_t^{stop} + \mu_2 n_t^{al,bike}) - \sum_{t'=2}^T \sum_{t=2}^{t'} n_{t-1}^{al,bike} n^u p \right] / (\varphi_1 V_D + \varphi_2 V_I) \quad (35)$$

where V_D and V_I are the coverage area for directed and undirected vertices, φ_1 and φ_2 are the weightage for each items.

5. Numerical example

5.1 Simplified Sioux Falls network

A typical benchmark network Sioux Falls network is applied to test the performance of the proposed algorithms, as shown in Figure 3. The network input data is extracted from Chakirov and Fourie (2014). The population data as listed in Table 2 is calculated based the household

data from Chakirov and Fourie (2014). This study assumes that the dockless bike sharing demand is 10% of the population at each vertex. All programs are run in MatlabR2017b on an Acer notebook with an Intel Core i5 CPU (1.70 GHz) under Windows 10 Home with 4 GB memory.

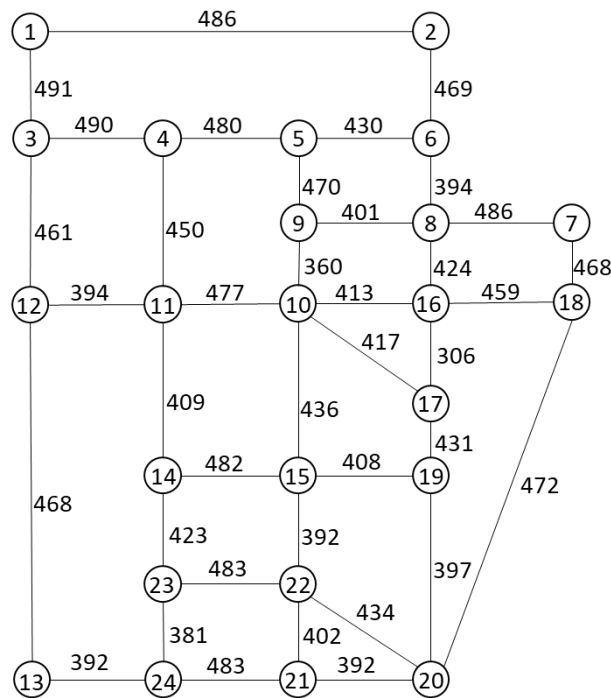


Figure 3 Simplified Sioux Falls network (the numbers on edges denote the distance between two connected vertices)

Table 2 Population data in study area

Vertex	1	2	3	4	5	6	7	8
Population	84	4362	3024	1685	2775	3422	3523	1956
Vertex	9	10	11	12	13	14	15	16
Population	3994	6248	5545	8332	9209	2324	5369	3337
Vertex	17	18	19	20	21	22	23	24
Population	2209	14783	5267	7085	2299	2738	2637	5283

Scenario A

Applying the algorithm in section 3.1, the result of modularity calculation is shown in Figure 4 and the dendrogram of the clustering result is shown Figure 5. Figure 4 shows the changes of modularity during clustering. When there are the 5 communities in the network, the value of modularity is greatest. Therefore, to achieve the market objective, 5 vertices need to be select in the network. Figure 5 shows how the vertices are distributed in each community. The numbers at the bottom of the dendrogram represent the vertices in the network. The dendrogram depicts how the vertices are merged together to form communities during the process at different levels. A cross-section of the dendrogram at any level gives the division of the subnetworks at that level. As obtained from Figure 4, $M = 0.465$ is the maximum modularity

corresponding to a configuration with 5 communities. Thus, the dendrogram needs to be cut for the split of 5 groups at the dash line depicted in Figure 5.

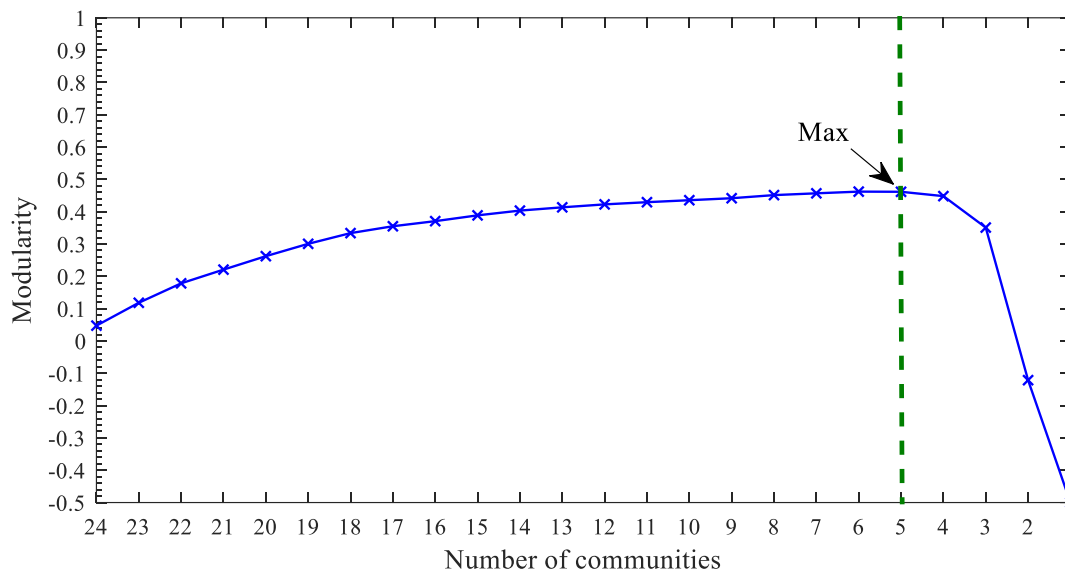


Figure 4 Modularity of partition results

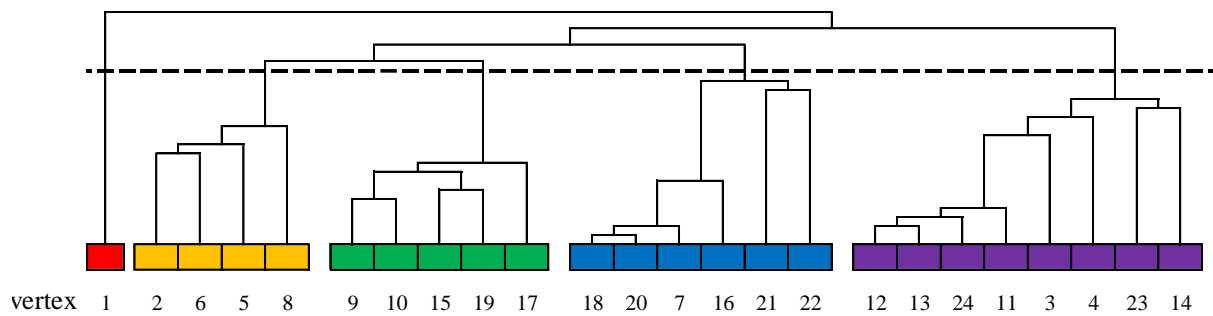


Figure 5 Dendrogram of clustering results

Figure 6 shows the detailed allocation position and coverage area based on the community structure. The optimal allocation position is the vertex with the greatest weight in each community. Specifically, vertex 1 itself is a single community with no connection with other vertices. Based on Table 1, it could be found that the demand level on vertex 1 is too small so that there is little trip from or to vertex. Hence, vertex 1 could be removed from the real allocation strategy.

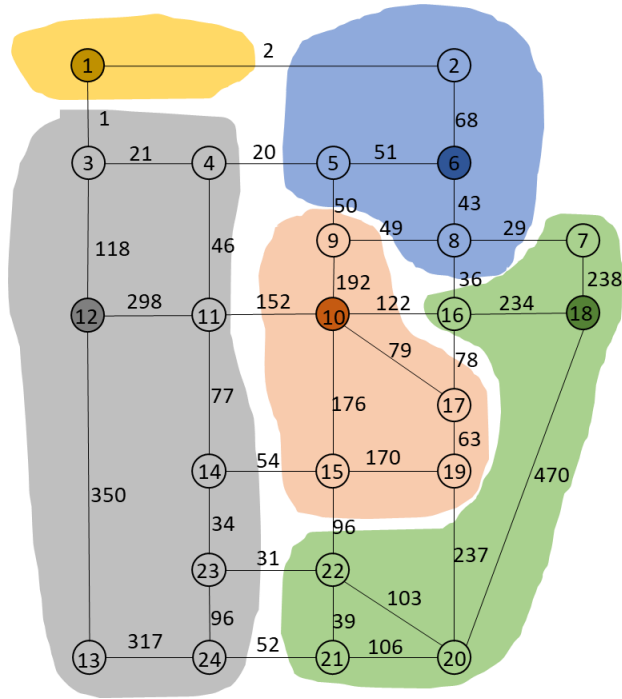


Figure 6 First allocation strategy in Scenario A (the number on the edge denotes edge weight)

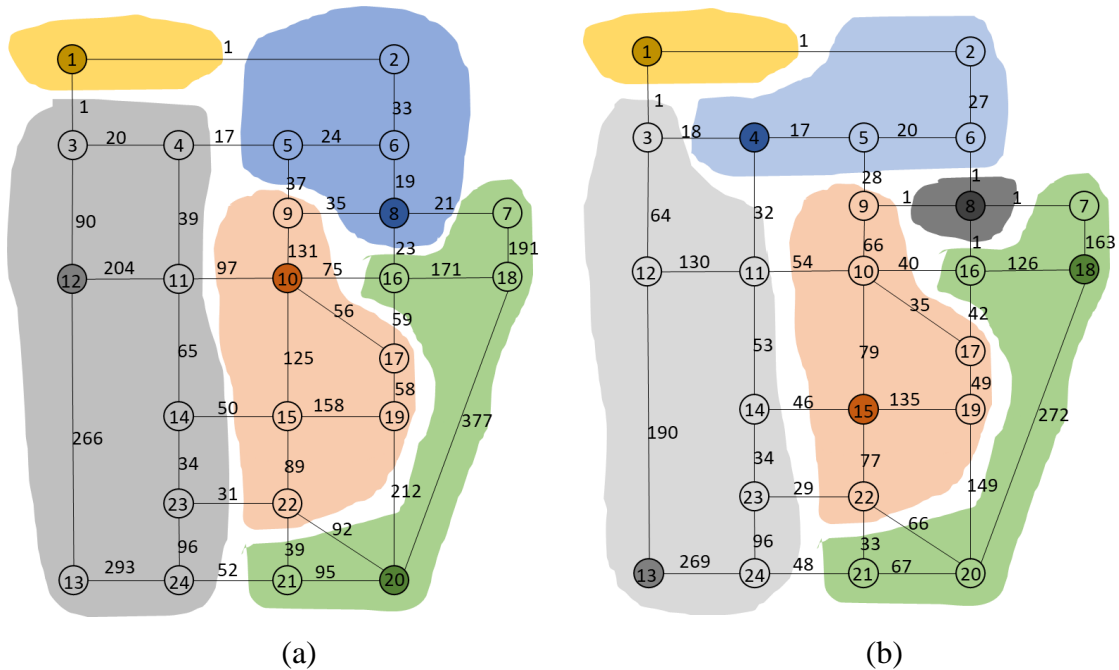


Figure 7 Second and third allocation strategies in Scenario A (the numbers on edge denotes the edge weight)

The process bikes allocation should be in batches from the time dimension no matter at the initial stage when the operator is trying to estimate the market demand or during operation when scrap bikes are away from the market. The usage could be obtained through GPS data in real application, or using Markov chain to simulate, while this study estimated the usage based on the weight of edge. Assume that 150 bikes have been allocated at the selected 4 vertices (6, 10, 12, 18), and each bike will be used twice per day. The number of trips on each edge is to multiply the total trips by the edge selection probability, which is based on the weight of edge. The updated demand is the difference between the total demand minus the number of trips. Applying the algorithm in section 3.1 again, the second allocation strategy is obtained as shown

in Figure 7(a). Following the same procedure, the third allocation strategy is shown in Figure 7(b). Compared Figure 7 with Figure 6, as the increase of the number of bikes in the market, the unsatisfied demand will lead to the changes of allocation position and coverage area. A new community structure with single vertex (8) is obtained. It means that the supply could meet the demand requirement at this vertex, which causes the less relationship of this vertex with other vertices.

Scenario B

Based on the diffusion theory, the allocation strategy should follow the same strategy obtained in Scenario A, which is to allocate the bikes at the vertex with the highest weight. Taken Figure 6 as an example, if the market follower selects other vertices than the market leaders to allocate the bikes (e.g. vertex 1), the market follower could quickly take the market share at vertex 1. However, as the demand at vertex 1 is relatively less, the total demand share of market follower is still at a greater disadvantage than the market leader.

Assume the current fare is 1 established by the market leader, there are total 100 rider demand in the market. There are already 50 riders that used the bikes from market leader. Among them, 20% may change to use the bikes from the market follower. Figure 8 depicts the relationship among market share of market follower, the ratio of the number of bikes from market follower to the total number of bikes in exiting market from market leader and the fare. Apparently, more bikes and lower price will enlarge the market share of market follower. If the market follower would like to achieve 40% of market share and use the current fare, 1.5 times of bikes than the market leader should be allocated at each position.

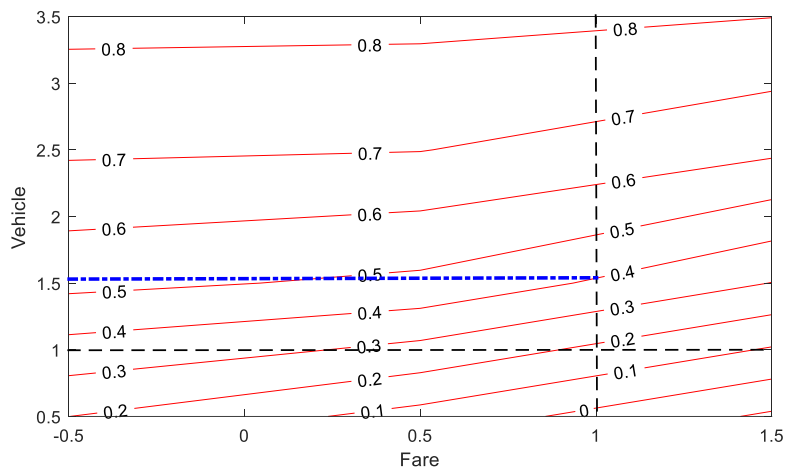


Figure 8 Relationship among market share, total number of bikes and fare for market follower

5.2 Large scale network in Singapore

As we have proved that the market follower should follow the same allocation strategy as market leader, market follower who wants to grasp the market should reduce the fare and increase the number of bikes. A real large-scale network in Singapore as shown in Figure 9 is used to test the practicability of the proposed allocation method. Assume the dockless bike sharing demand is in direct proportion to the population size in each TAZ, and the population size in each TAZ is in direct proportion to the maximum plot ratio. The maximum plot ratio is equal to the maximum permissible plot ratio marked in the master plan times the area of TAZ. The distance matrix is extracted from google map by using the minimum walking distance. The running time is 0.037 second.

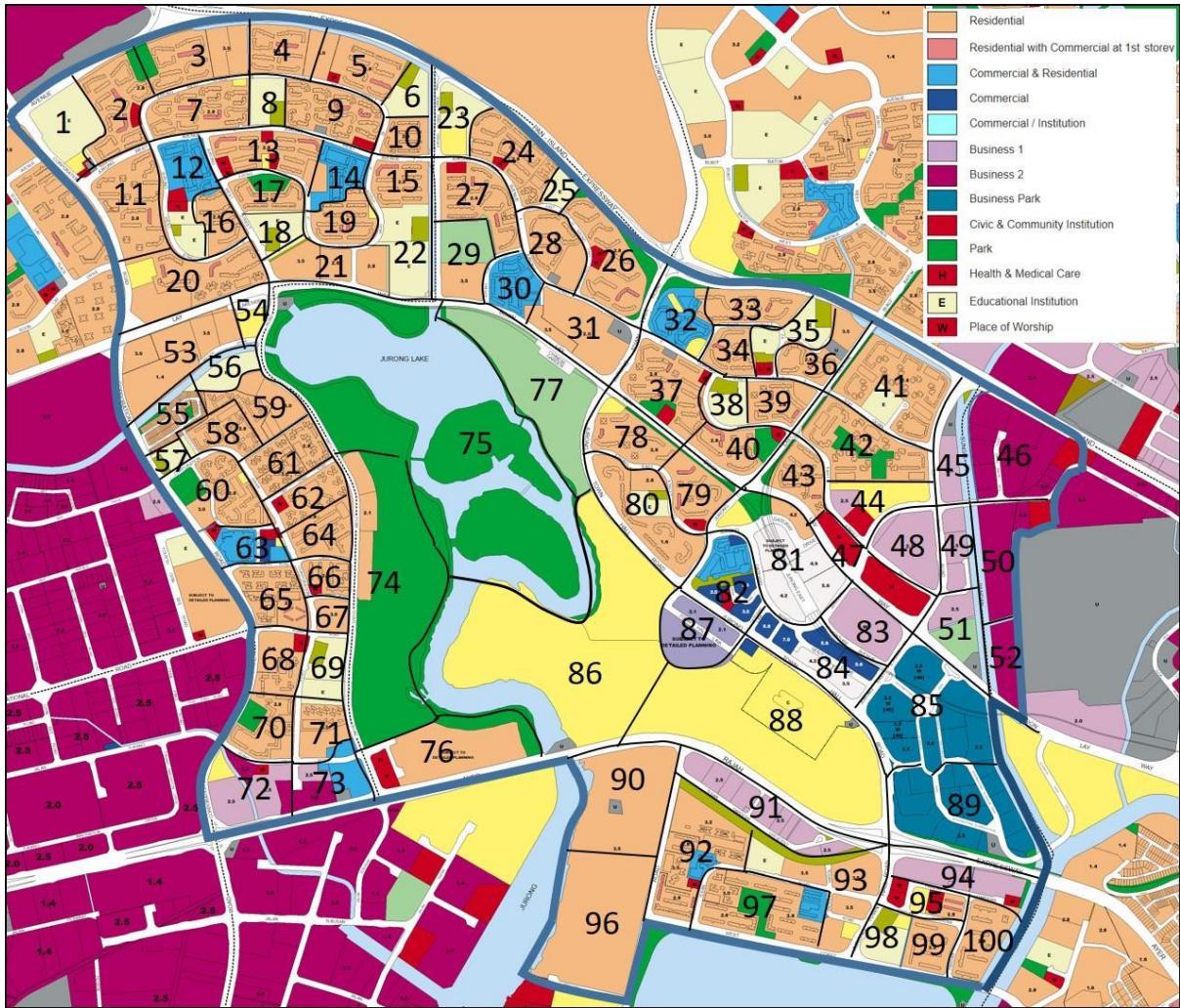


Figure 9 Study area in Jurong Lake District in Singapore
<https://www.ura.gov.sg/maps2/?service=mp#>

Figure 10 shows the first allocation position and service area in the study area. Compared with Figure 9, it could be found that lines 1 to 3 are all rivers, which increase the trip distance between vertices on both side of the river. The long trip distance decreases the connections between vertices on both side of the river and thus forms the boundaries of the community structure. Although the express way of line 4 increase the trip distance between vertices on both side, the vertices are still in the same community. It is due to the vertices on the upper side of the express way are mostly in business or commercial area. Riders could not cross the building like in residential area to the other side to seek for the shortest distance. Riders from vertices like 52, 85 89 to vertex 81 must follow the road so that the real trip distance is larger than the linear distance between coordinates. Metro stations (e.g. vertices 54, 77 and 81), shopping mall (e.g. vertices 54, 77 and 81) and other vertices with high plot ratio (e.g. vertices 58 and 92) have all been selected as allocation positions, which is reasonable in real. Because they are with high demand level and play the crucial role in the network.

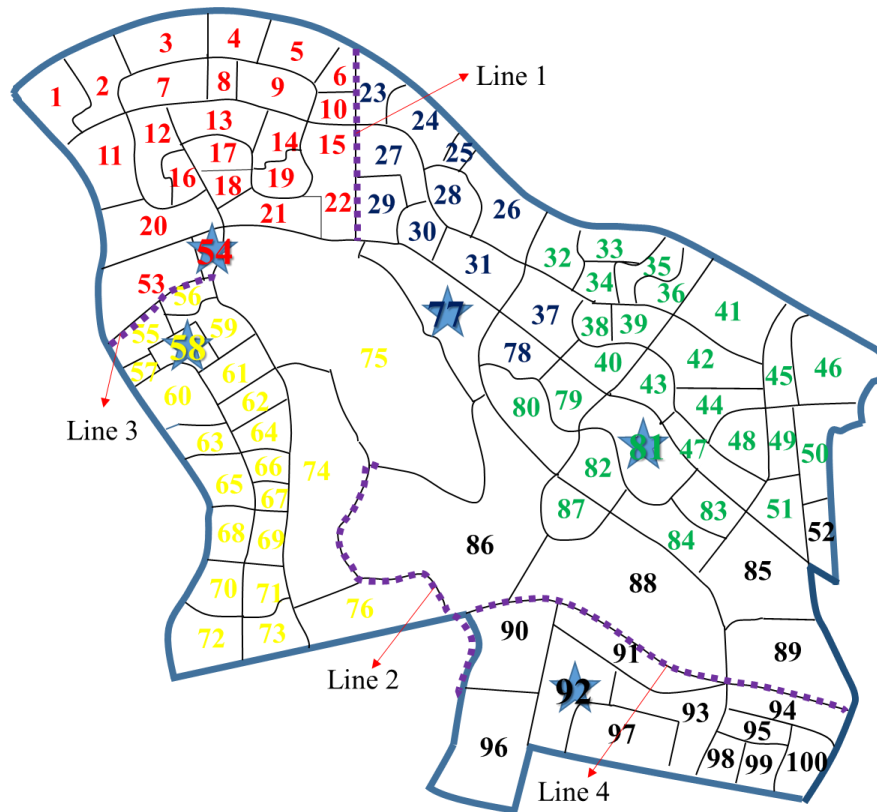


Figure 10 Allocation strategy using the proposed method

As there is no existing method that can solve the dockless bike sharing allocation problem from literature, we compare our results with the results from a Cellular Automata based station selection study (Meng et al., 2015) in same network. In station-based bike sharing system, a station is determined mainly due to the high demand, which follows the same principle as dockless bike sharing allocation strategy making. Results in Figure 11 has been validated by the local planning authority in the station-based bike sharing system plan. Comparing with the results in Figure 10, all the allocation locations have been covered in Figure 11, which demonstrated the accuracy of the proposed method. In station-based bike sharing system, stations will be planned more than the allocation locations in dockless bike sharing system. Because the station-based bike sharing system considers the future demand for development while the dockless bike sharing allocation aims to meet the current requirements.

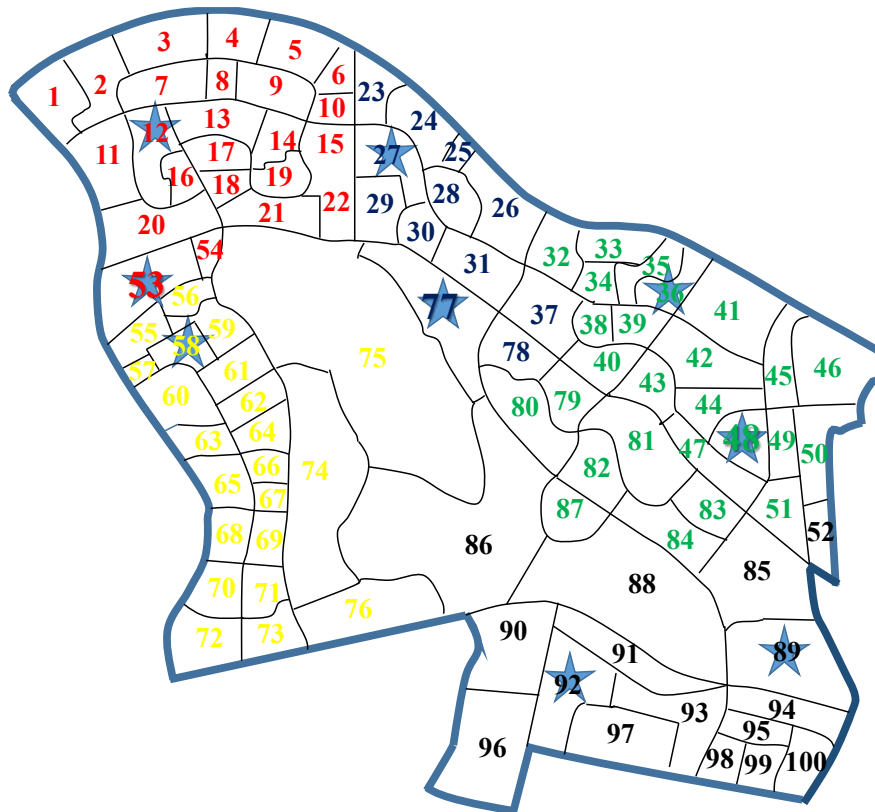


Figure 11 Station location planning in the study area

6. Conclusions

This paper proposed a framework to guide the bike allocation strategy in a competitive dockless bike sharing market. Scenarios for both market leader and follower have been considered. An improved community structure detection algorithm is developed to determine the boundaries of the service area for each allocation position. Moreover, a systemic bi-level approximate optimal allocation model is proposed. For market leader, as the demand level is not clear at the beginning, the allocation should be made in small batches. The vertices with maximum weight shall be used to allocate bikes. This strategy could distribute the bikes with the fastest speed. For market follower, same allocation strategy should be adopted based on diffusion theory. Market follower should use lower price, more bikes, and large batches allocation to strive for market share.

The work could be extended future by: (1) optimising the number of bikes in each allocation considering the budget limitation; (2) investigating the operation strategies in a mixed station-based and dockless bike sharing system; (3) discussing the optimal pricing strategies for both market leader and follower in the competitive market.

Acknowledgement

This research is partly supported by the Singapore Ministry of Education (MOE) AcRF Tier 2 Grant MOE2016-T2-1-044.

Reference

Bordagaray, M., dell'Olio, L., Fonzone, A., & Ibeas, Á. (2016). Capturing the conditions that introduce systematic variation in bike-sharing travel behavior using data mining techniques. *Transportation Research Part C: Emerging Technologies*, 71, 231-248.

- Chang, Y. T., Leahy, R. M., & Pantazis, D. (2012). Modularity-based graph partitioning using conditional expected models. *Physical Review E*, 85(1), 016109.
- Chakirov, A., & Fourie, P. J. (2014). Enriched sioux falls scenario with dynamic and disaggregate demand. *Arbeitsberichte Verkehrs-und Raumplanung*, 978.
- Chen, Q., Liu, M., & Liu, X. (2018). Bike fleet allocation models for repositioning in bike-sharing systems. *IEEE Intelligent Transportation Systems Magazine*, 10(1), 19-29.
- Datner, S., Raviv, T., Tzur, M., & Chemla, D. (2017). Setting inventory levels in a bike sharing network. *Transportation Science*, 53(1), 62-76.
- Faghih-Imani, A., Hampshire, R., Marla, L., & Eluru, N. (2017). An empirical analysis of bike sharing usage and rebalancing: Evidence from Barcelona and Seville. *Transportation Research Part A: Policy and Practice*, 97, 177-191.
- Fishman, E. (2016). Bikeshare: A review of recent literature. *Transport Reviews*, 36(1), 92-113.
- Forma, I. A., Raviv, T., & Tzur, M. (2015). A 3-step math heuristic for the static repositioning problem in bike-sharing systems. *Transportation Research Part B: Methodological*, 71, 230-247.
- Frade, I., & Ribeiro, A. (2015). Bike-sharing stations: A maximal covering location approach. *Transportation Research Part A: Policy and Practice*, 82, 216-227.
- García-Palomares, J. C., Gutiérrez, J., & Latorre, M. (2012). Optimizing the location of stations in bike-sharing programs: A GIS approach. *Applied Geography*, 35(1-2), 235-246.
- Girvan, M., & Newman, M. E. (2002). Community structure in social and biological networks. *Proceedings of the national academy of sciences*, 99(12), 7821-7826.
- Ho, S. C., & Szeto, W. Y. (2014). Solving a static repositioning problem in bike-sharing systems using iterated tabu search. *Transportation Research Part E: Logistics and Transportation Review*, 69, 180-198.
- Ho, S. C., & Szeto, W. Y. (2017). A hybrid large neighborhood search for the static multi-vehicle bike-repositioning problem. *Transportation Research Part B: Methodological*, 95, 340-363.
- Kabak, M., Erbaş, M., Çetinkaya, C., & Özceylan, E. (2018). A GIS-based MCDM approach for the evaluation of bike-share stations. *Journal of Cleaner Production*, 201, 49-60.
- Hsu, C. C., Liou, J. J., Lo, H. W., & Wang, Y. C. (2018). Using a hybrid method for evaluating and improving the service quality of public bike-sharing systems. *Journal of Cleaner Production*, 202, 1131-1144.
- Liu, J., Li, Q., Qu, M., Chen, W., Yang, J., Xiong, H., ... & Fu, Y. (2015, November). Station site optimization in bike sharing systems. In *2015 IEEE International Conference on Data Mining* (pp. 883-888). IEEE.
- Malliaros, F. D., & Vazirgiannis, M. (2013). Clustering and community detection in directed networks: A survey. *Physics Reports*, 533(4), 95-142.
- Meng, M., Zhang, J., & Wong, Y.D. (2015). A GIS-based Cellular Automata model for planning bike-sharing network. In *the proceeding of mobil.TUM 2015*. Munich, Germany.
- Meng, M., Zhang, J., Wong, Y. D., & Au, P. H. (2016). Effect of weather conditions and weather forecast on cycling travel behavior in Singapore. *International journal of sustainable transportation*, 10(9), 773-780.

- Newman, M. E. (2004). Fast algorithm for detecting community structure in networks. *Physical review E*, 69(6), 066133.
- Ofo. (2018). The world leader in station-free bike sharing. Access from: <https://www.ofo.com/sg/en/about>
- Pal, A., & Zhang, Y. (2017). Free-floating bike sharing: solving real-life large-scale static rebalancing problems. *Transportation Research Part C: Emerging Technologies*, 80, 92-116.
- Pan, L., Cai, Q., Fang, Z., Tang, P., & Huang, L. (2018). Rebalancing dockless bike sharing systems. *arXiv preprint arXiv:1802.04592*.
- Preece, J., & Shneiderman, B. (2009). The reader-to-leader framework: Motivating technology-mediated social participation. *AIS transactions on human-computer interaction*, 1(1), 13-32.
- Raviv, T., & Kolka, O. (2013). Optimal inventory management of a bike-sharing station. *Iie Transactions*, 45(10), 1077-1093.
- Shen, Y., Zhang, X., & Zhao, J. (2018). Understanding the usage of dockless bike sharing in Singapore. *International Journal of Sustainable Transportation*, 1-15.
- Shu, S., Bian, Y., Rong, J., & Xu, D. (2019). Determining the exact location of a public bicycle station—The optimal distance between the building entrance/exit and the station. *PLoS one*, 14(2), e0212478.
- Szeto, W. Y., Liu, Y., & Ho, S. C. (2016). Chemical reaction optimization for solving a static bike repositioning problem. *Transportation Research Part D: Transport and Environment*, 47, 104-135.
- Wu, Q., Qi, X., Fuller, E., & Zhang, C. Q. (2013). “follow the leader”: A centrality guided clustering and its application to social network analysis. *The Scientific World Journal*, 2013. ID 368568.
- Yan, K. K., Lou, S., & Gerstein, M. (2017). MrTADFinder: A network modularity based approach to identify topologically associating domains in multiple resolutions. *PLoS Computational Biology*, 13(7), e1005647.
- Yang, X. H., Chen, G., Chen, S. Y., Wang, W. L., & Wang, L. (2014). Study on some bus transport networks in China with considering spatial characteristics. *Transportation Research Part A: Policy and Practice*, 69, 1-10.
- Zhang, L., Zhang, J., Duan, Z. Y., & Bryde, D. (2015). Sustainable bike-sharing systems: characteristics and commonalities across cases in urban China. *Journal of Cleaner Production*, 97, 124-133.
- Zhang, J., Koh, P. P., Meng, M., Leow, B. W., & Wong, Y. D. (2017). Feasibility study of bike sharing in Singapore. *Proceedings of the Institution of Civil Engineers-Engineering Sustainability*. 171(7), 351-357.
- Zhang, Y., Lin, D., & Mi, Z. (2019). Electric fence planning for dockless bike-sharing services. *Journal of Cleaner Production*, 206, 383-393.
- Zhou, M. Y., Cai, S. M., & Fu, Z. Q. (2012). Traffic dynamics in scale-free networks with tunable strength of community structure. *Physica A: Statistical Mechanics and its Applications*, 391(4), 1887-1893.

Zhou, S., & Ni, Y. (2018). Effects of dockless bike on modal shift in metro commuting: A pilot study in Shanghai. *Transportation Research Board 97th Annual Meeting*, Washington DC, United States. (No. 18-04973).