Towards Effective Top-k Location Recommendation for Business Facility Placement

Pu Wang, Wei Chen, and Lei Zhao(✉)

School of Computer Science and Technology,
Soochow University, Suzhou 215006, China
{wangpu,robertchen,zhaol}@suda.edu.cn

Abstract. In the process of selecting locations for establishing new business facilities, location recommendation offers the optimal candidates, which maximizes the number of customers served to bring the maximum profits. In most existing work, only spatial positions of customers are considered, where social relationships and temporary activities, which are significant factors for candidate locations, are ignored. Additionally, current studies fail to take the capacity of service facilities into consideration. To overcome the drawbacks of them, we introduce a novel model MITLR (Multi-characteristic Information based Top-k Location Recommendation) to recommend locations with respect to capacity constraints. The model captures the spatio-temporal behaviors of customers based on historical trajectory and employs social relationships simultaneously, to determine the optimal candidate locations. Subsequently, by taking advantage of feature evaluating and parameter learning, MITLR is implemented through a hybrid B-tree-liked framework called CLTC-forest (tree). Finally, the extensive experiments conducted on real-world datasets demonstrate the better effectiveness of proposed MITLR.

Keywords: Top-k location recommendation · Spatio-temporal trajectory · Social relationship · Capacity constraint

1 Introduction

The study of top-k facility locations selection aims to identify the appropriate k locations for new business facilities from a range of available candidate locations. In this context, location is defined as a special site in a road network with facility on it in terms of a given service for customers, the selection is based on factors such as the number of customers served or the returns on facility investments. This kind of query has been widely applied in a variety of recommendation applications, such as planing to establish new electric vehicle charging stations, mobile toilets, or retail stores in a city.

As GPS and mobile devices are developed in recent years, daily trajectories have been recorded and utilized widely, as well as an increasing number of

G. Li et al. (Eds.): KSEM 2020, LNAI 12275, pp. 51–63, 2020.
https://doi.org/10.1007/978-3-030-55393-7_5

studies with location recommendation come to focus on trajectories [1,4–6,10]. Furthermore, advances in social network technology are facilitating interpersonal communication, friend-based recommendation becomes a growingly significant and relevant factor in recommendation system. For instance, customers will receive electronic red envelopes or coupons occasionally while consuming, and then share them with friends in some instant messaging Apps like Wechat, QQ, Alipay, etc. With shared electronic red envelopes or coupons, their friends could get a discount when they are consuming afterwards [11].

However, existing studies of trajectory-based location recommendation evaluate the correlation between customers and facility locations by their spatial distances solely, they fail to evaluate the timestamp of trajectories and the effects of customer social relationships on facility locations querying, which will render the recommendation results inaccurate or uneconomic. To illustrate the necessity of considering the friend-based recommendation, a straightforward example is demonstrated as below.

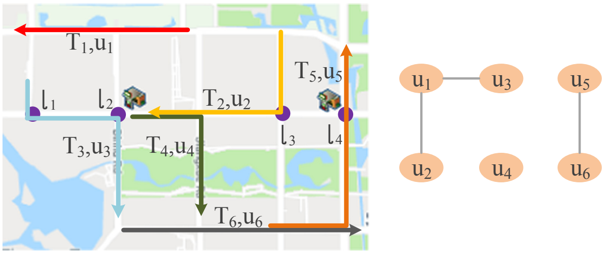

Fig. 1. A toy example with candidate locations, trajectories and social relationships.

Example 1. As shown in Fig. 1, there are four candidate locations l_1 to l_4 with electric vehicle charging stations, and six vehicle trajectories T_1 to T_6 corresponding to six distinct customers u_1 to u_6. To select the optimal facility location for u_6, as the shortest spatial distances from T_6 to l_2, l_3 and l_4 are all equal, a random one between l_2, l_3 and l_4 would be chosen by several existing algorithms such as NetClus [6], while spatial distances with candidate locations are considered merely. However, Fig. 1 shows that u_5 and u_6 could readily share electronic red envelopes or coupons since they are close friends. Therefore, if u_5 has been served by l_4 (as u_5 passes l_4 directly), there is a great probability that u_6 will also be served by l_4 due to the reciprocal recommendation of u_5.

To overcome the aforementioned deficiencies of earlier work, we formalize the problem of constrained top-k facility location recommendation into a novel model MITLR. To determine whether a candidate location is recommended or not, the total service utilities of candidate location are predicted. Unlike the previous work that only considers the road spatial distance, this model examines the importance of spatial and temporal features of trajectories, as well as the

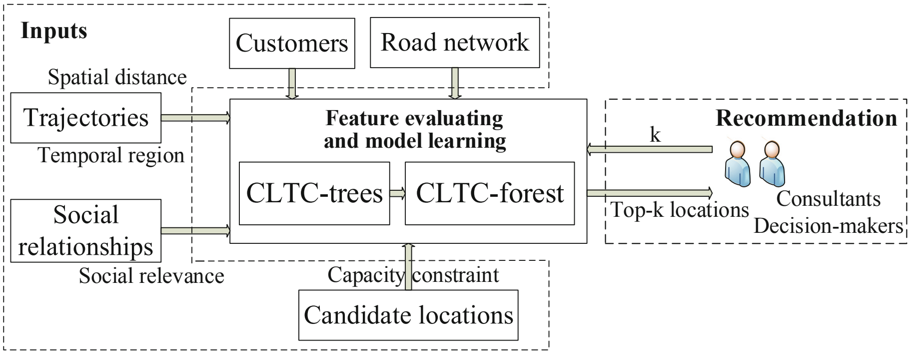

Fig. 2. The framework of MITLR.

social relationships of customers at the same time. As a result, the evaluation of service utilities is rendered more precisely by these significant characteristics and the new objective function. In addition, we take a formal approach to consider the capacity of all candidate locations with regards to the real living conditions of customers. We then develop a new hybrid index using service utilities and location capacities, which is referred to CLTC (Constrained Location and social-Trajectory Clusters) forest, to incorporate spatio-temporal trajectories, social relationships and candidate locations into CLTC-trees that ultimately form the final CLTC-forest. Based on this hybrid index, this study presents an efficient query algorithm that exploits a simple greedy strategy to obtain the final top-k results for recommending.

The overall framework of the proposed model is outlined in Fig. 2, and our key contributions are summarized as follows.

1. This study proposes the top-k location recommendation problem and a novel model MITLR, and defines the service utility function to predict the correlation while capturing the spatio-social and temporal behaviors.
2. We have developed a new index structure of the CLTC-forest (tree) which combines both candidate locations and trajectories, as well as present the process of parameter learning and efficient query approach.
3. Extensive experiments are performed on real datasets to offer insight into the effectiveness and feature weights of the proposed model.

2 Related Work

We cover the existing work on next optimal location prediction and the optimal k locations prediction in turn.

Several studies [1–3,7,9] focus on the next optimal location prediction problem by taking advantage of various metrics and objective functions. Sun et al. [7] acknowledge the service capacity of each location facility, they suggest that there is a certain limitation on the number of served customers and the metric is only examined by spatial distances. Li et al. [3] query a location for establishing facility based the optimal segment query of trajectories, they assign a score on each segment as in [1] but without recognizing the candidate location as a specific position on a road network. Yao et al. [9] take advantage of a recurrent model SERM for the next optimal location prediction in semantic trajectories, where both spatio-temporal transitions and relevant textual information posted by customers are considered to improve the precision.

Recent researches have concentrated on exploring the problem of the optimal k location recommendation [4–6,10]. In more details, Li et al. [4] mine the most influential k-location, from this point of view, they evaluate the maximum number of unique trajectories that traverse a location in a given spatial region, therefore, the common practicability of this work is greatly restricted by the traverse limitation. Mitra et al. [5,6] focus on the top-k location query problem with respect to trajectory merely, they [6] propose an index framework of NetClus for TOPS query by covering a wide range of objective functions, their work assumes that each of candidate locations has a radius parameter τ, as a result, the construction of NetClus leads to lots of index instances which are calculated and stored with different values of τ and cluster radii. They further extend TOPS to TIPS [5] in an attempt to minimize either the maximum customer inconvenience or the average inconvenience.

3 Problem Statement

In this section, we formalize the problem and the model MITLR, Table 1 summarizes the frequently used notations throughout this paper.

Table 1. Notation and corresponding description.

Notation	Description
L, C	Set of candidate locations and capacities
Γ	Set of spatio-temporal trajectories
U, S	Set of customers and social relationships
$f_d(l_i, u_j)$	Spatial distance of l_i and u_j
$f_s(l_i, u_j)$	Social relevance of l_i and u_j
$f_t(l_i, u_j)$	Temporal region of l_i and u_j
$F(u_j)$	Friend set of u_j
$SU(l_i, u_j)$	Service utility of l_i and u_j
$\Psi(l_i)$	Service utilities of l_i
$\Im$	Set of optimal k locations

Considering a setting where candidate locations and trajectories are located in a road network and social relationships of customers in trajectories can be captured. The road network is defined as a directed weighted graph $G = \{V_g, E_g\}$, where V_g denotes the set of vertices of road intersections and E_g denotes the set of directed edges of road segments, the weight of directed edge denotes its spatial distance. The candidate location l is a place for establishing a certain facility or service like electric vehicle charging station or mobile toilet, and the service capacity of l_i is defined as c_{l_i}, which means that it cannot be exceeded in real serving applications. A trajectory T is represented in the sequential form of $T = \{(v_1, t_1), ..., (v_\zeta, t_\zeta)\}, v_\zeta \in V_g$, where t_ζ denotes the timestamp when T crosses v_ζ. The social relationships of customers are simply modeled as an undirected and unweighted graph $S = \{V_u, E_u\}$, V_u is the set of nodes representing the customers, and E_u is the set of edges, where an edge denotes that there is a friend relationship between two corresponding customers. Besides, we suppose $|L| = m$ and $|\Gamma| = |U| = n$.

First of all, three significant characteristics of spatial distance, social relevance and temporal region in service utility are provided in detail.

Spatial Distance. The formula of the shortest spatial distance between location and trajectory is adopted, as presented in [6]. Therefore, the spatial distance characteristic of customer and location is outlined:

$$f_d(l_i, u_j) = \frac{\min_{\forall v_{jk}, v_{j\iota} \in T_j} \{d(v_{jk}, l_i) + d(l_i, v_{j\iota}) - d_a(v_{jk}, v_{j\iota})\} - \rho_{\min}}{\rho_{\max} - \rho_{\min}} \tag{1}$$

where $\rho_{\min}$ and $\rho_{\max}$ are normalization factors, $d_a(v_{jk}, v_{j\iota})$ denotes the shortest spatial distance from v_{jk} to $v_{j\iota}$ by going along T_j. To illustrate, a customer u_j deviates from v_{jk} of T_j to l_i, and then returns to $v_{j\iota}$ on G in her/his usual trajectory, note that the additional distance on T_j is not included.

Social Relevance. For assessing the social relevance of l_i and u_j, we assume that $F(u_j)$ is set of customers who have friend relationships with u_j, and $CU(l_i)$ represents set of customers that have already been evaluated to l_i for being served. Then the social relevance between u_j and l_i is defined as follows:

$$f_s(l_i, u_j) = \frac{|\{u_k | u_k \in F(u_j) \wedge u_k \in CU(l_i)\}|}{|CU(l_i)| + \lambda_s} \tag{2}$$

where $|\cdot|$ denotes the number of elements and λ_s is the Laplace smoothing coefficient. The intuition behind the social relevance feature is the friend-based recommendation through shared electronic red envelopes or coupons, and only the direct friendships between customers are concerned here.

Temporal Region. Supposing that u_j departs from her/his usual trajectory to one location at timestamp t_{jk} in Eq. (1), and arrives at location l_i at timestamp $AT(u_j, l_i) = t_{jk} + \Delta t$, where Δt is a constant timestamp value representing the duration from departure to arrival. For simplicity, Δt of each customer is set to equal and one day is divided into 24 equal segments. As a consequence, the temporal region characteristic is given:

$$f_t(l_i, u_j) = \frac{\sum\limits_{u_k \in CU(l_i)} \min\{|\mathrm{I}(u_j, u_k, l_i)|, 24 - |\mathrm{I}(u_j, u_k, l_i)|\}/24}{|CU(l_i)| + \lambda_t} \tag{3}$$

where $\mathrm{I}(u_j, u_k, l_i) = AT(u_k, l_i) - AT(u_j, l_i)$, and λ_t is also the Laplace smoothing coefficient. The ground truth of temporal region is that, if a customer intends to stagger her/his arrival time with others who have already been evaluated to be served in the same facility, then the customer will get more guaranteed service.

Next, the service utility function that is raised to evaluate the correlation between l_i and u_j can be presented:

$$SU(l_i, u_j) = \begin{cases} 1 & f_d(l_i, u_j) = 0 \\ -\alpha_\vartheta * f_d(l_i, u_j) + \beta_\vartheta * f_s(l_i, u_j) + \gamma_\vartheta * f_t(l_i, u_j) & otherwise \end{cases} \tag{4}$$

where α_ϑ, β_ϑ and γ_ϑ are feature weights and $\alpha_\vartheta + \beta_\vartheta + \gamma_\vartheta = 1$, if $f_d(l_i, u_j) = 0$, it demonstrates that T_j of u_j just traverses l_i straightforward. The greater value of $SU(l_i, u_j)$ indicates there is a closer connection between l_i and u_j, and there is also a higher probability of l_i serving u_j, vice versa. Moreover, from the perspective of a candidate location, the total service utilities (also as service revenue) for all the served customers are defined as follows:

$$\Psi(l_i) = \sum_{u_j \in U'} SU(l_i, u_j), U' \in U, |U'| \leq C_i \tag{5}$$

where $\forall u_v \in U', \forall u_l \in U - U', SU(l_i, u_v) \geq SU(l_i, u_l)$. Consequently, the model of MITLR is formally stated.

Problem Definition. Given a query with parameter k, a set of spatio-temporal trajectories Γ with corresponding customers U on G, a set of social friend relationships S, and a set of candidate locations L with capacities C, the MITLR seeks to select the optimal location set $\Im$ ($\Im \in L, |\Im| = k$), which maximizes the sum of total service utilities Υ without exceeding the capacity limitations of each selected location, where $\Upsilon = \arg\max \sum_{i=1}^{k} \Psi(l_i), l_i \in \Im$.

4 Model Implementation

It is recognized that candidate locations in close proximity are prone to serve a great number of identical customers, and when $k \ll m$ and $c_i \ll n, i \in [1, m]$, the locations selected in query results are all keeping a certain distance from each other, therefore, the CLTC-tree (forest) is designed for MITLR in terms of model implementing and model learning. For specified k, the final k locations are returned from CLTC-forest by adopting a simple greedy manner.

4.1 CLTC-forest

The CLTC-forest is composed of a series of CLTC-trees, and the CLTC-tree is a B-tree-liked hybrid index structure that integrates candidate locations, trajectories, and customers with social relationships according to their service utilities and location capacities. Each tree node links three pieces of additional information, which include a candidate location l_i as its representation (label), a set of

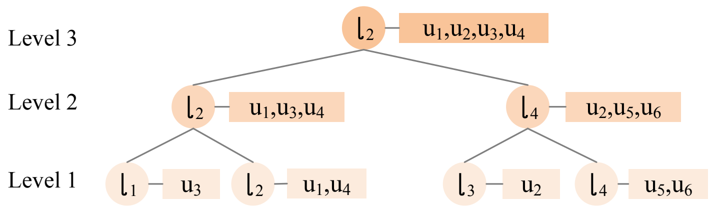

Fig. 3. A simple CLTC-tree.

customers that will be served by $U(l_i)$ with their corresponding SU limited by facility capacity, as well as a total service utility of Ψ.

However, the none-leaf nodes of CLTC-tree are quite different from leaf nodes, as the representation of a none-leaf node is one of the labels of its two children nodes, the customers are a subset of a union which is constituted by customers of two children nodes that do not exceed the capacities. A simple example of one CLTC-tree that represents candidate locations, trajectories and customers with corresponding social relationships of Fig. 1 is denoted as in Fig. 3, where Ψ is omitted in each of specific tree node.

4.2 Generation Processes

The CLTC-forest construction includes two steps of clustering and combining.

Clustering. In order to cluster candidate locations with their most relevant customers into leaf nodes, a constrained k-medoid-liked cluster algorithm is proposed, in which all of candidate locations are served as 'medoids' and service utilities calculated by Eq. (4) are served as the metric. It can be seen that the final clusters are constrained by service capacities of each location on the scales.

However, if a number of customers that pass one candidate location simultaneously happen to overwhelm the capacity of the location, those customers whose initiating position or ending position is in the proximity of this location are clustered firstly. The is because that a customer will not be visibly disturbed or interrupted if she/he chooses one facility near to the initiating position or the ending position on the trace of her/his trajectory.

Combining. Since each leaf node is also a simple CLTC-tree which has merely one root node, a series of merging approaches can be adopted to combine two CLTC-trees into one while utilizing their root node information. By repeating this process, the entire CLTC-trees from leaf nodes to root nodes could be constructed, and then, CLTC-forest is formed finally. Before introducing the CLTC-trees-merging, the definition of the coherence of two CLTC-trees Ct_i and Ct_j is proposed as $CO(Ct_i, Ct_j) = \max\{\Psi(Lrn(Ct_i)), \Psi(Lrn(Ct_j))\}$, where $Lrn(Ct_i)$ is the representation of root node in Ct_i. $CO(Ct_i, Ct_j)$ indicates the service utilities that combines Ct_i and Ct_j into one by taking advantage of the additional information of two separate former root nodes. Greater value of $CO(Ct_i, Ct_j)$

represents that there is closer relevance of the two CLTC-trees, such as intimate relationships between the two customers set, or adjacent spatial distances between candidate locations. Therefore, CLTC-trees can be combined according to their coherence.

If two CLTC-trees is merged into one, a new tree node will be created to represent the root node, where two sub-trees are the two former CLTC-trees. For the newly root node, the service utility is equal to CO, the set of evaluated customers is U', where $U' \in Urn(Ct_i) \cup Urn(Ct_j)$ by Eq. (5), Urn is the corresponding customers in tree node, and its representation is the candidate location with larger value of CO in two children CLTC-trees.

4.3 Model Learning

With the help of Eq. (4), it can be observed that the service utility is just referred as a linear combination of the inputs, accordingly, several different kinds of regression algorithms could be deployed to learn these parameters. In this study, a linear regression with regularization is utilized, which the goal is to minimize the error e_M between the ground-truth location facilities and the recommended location results returned. Supposing that the overall parameters are denoted as $\theta_M(\alpha_\vartheta, \beta_\vartheta, \gamma_\vartheta)$, then the corresponding optimization function is defined as:

$$\min_{\theta_M} \sum\nolimits_{i=1}^{k} {d_E}^2(l_{pi}, l_{ri}) + \gamma_e ||\theta_M||^2 \tag{6}$$

where γ_e is the regularization parameter and set equal to 10^{-8} as demonstrated in [2], l_{pi} is the predicted location and l_{ri} is the corresponding ground truth location, $d_E(l_{pi}, l_{ri})$ indicates the Euclidean distance between l_{pi} and l_{ri}.

4.4 Location Recommending

In querying, the tree levels of CLTC-forest are marked in a top-down fashion firstly, we and assume that the highest tree level is $\hbar$, where $\hbar < \lg(\lfloor m \rfloor + 1)$. All of the root nodes of CLTC-trees are marked with $ls_\hbar$, and the children of root nodes (in $ls_\hbar$, if have) are marked as $ls_{\hbar-1}$, by repeating the process until there is no node left to be marked (until to ls_1), then the entire tree level marks with the corresponding tree nodes are inserted into a set LS. Within the specified k, the mark of tree level is selected while $|ls_i| = k$, if exists, the k candidate locations in the nodes of ls_i are the querying results. However, if the mark does not exist, two marks of ls_i and ls_{i-1} are chosen where $|ls_i| < k < |ls_{i-1}|$. Subsequently, by utilizing a simple greedy manner, the k distinct candidate locations, which boast the maximum total service utilities, are selected to the recommendation results from the nodes of ls_i and ls_{i-1}. It is noticed that node in ls_i and two of its children nodes in ls_{i-1} in one CLTC-tree share one label, so they could not be chosen into the result together.

5 Experimental Evaluation

5.1 Datasets

The most widely used urban datasets of Beijing and Shanghai are employed in this study, where the intersections of road network are utilized to represent candidate locations. To simulate and generate customers with the corresponding trajectories, algorithm of discovering *Popular Routes* in [8] has been adopted with two real urban datasets. We extract the customer check-in and following data from Sina WeiBo[1], where there is a timestamp in each check-in point that is accurate to seconds, as well as a part of the trajectory traces of automobiles are collected. The social relationships are also extracted from Sina WeiBo, where two customers who are following with each other show that they are close friends. The statistics of datasets are listed in Table 2.

Table 2. Statistics of the datasets.

Categories	Beijing	Shanghai
# of intersections	171,186	333,766
# of road segments	226,237	440,922
# of customers (trajectories)	412,032	230,303
# of candidate locations	171,186	333,766
# of social relationships	26,139,861	13,687,459
# of CR	300	313
# of VCS	854	868

Two categories of popular existed facilities are prepared for model training and model testing in two cities, which are chained restaurants (CR) including KFC, MacDonald and Pazza Hut, as well as fast vehicle charging stations (VCS). The number of two existed facilities is also presented in Table 2 respectively. Furthermore, their geographic coordinates are obtained from AutoNavi[2].

5.2 Evaluation Plans

Competitive Approaches. To the best of our knowledge, no existing studies have been committed to the top-k candidate recommendation by exploiting customer trajectories and social relationships in a city-scale road network so far, therefore, we compare our model with a series of competitive methods, which are k-Medoids, SERM [9] and NetClus [6] by means of slight modifications.

Evaluation Metrics. A couple of the metrics of *Precision* and *Root Mean Square Error* (RMSE) are designed carefully in effectiveness evaluating. On one

[1] https://www.weibo.com/.
[2] https://www.amap.com/.

hand, we suppose that $\Im$ is the top-k querying results obtained from testing data, as well as L' is the corresponding existed facilities with the same category, then the precision is given as $\mathrm{P}_k = \frac{\sum_{i=1}^{k} hit(\Im_i, L'_i)}{k}$, where $hit(\Im_i, L'_i) = 1$ indicates there is a corresponding facility L_i^{ϑ} that satisfies $d_E(\Im_i, L'_i) \leq \tau_E, \Im_i \in \Im$ and $L'_i \in L^{\vartheta}$. On the other hand, RMSE is also adopted to measure the deviations between the recommended locations and the ground-truth facilities, the definition of RMSE is $RMSE_k = \sqrt{\frac{\sum_{i=1}^{k} \min\left(d_E(\Im_i, L'_i)\right)^2}{k}}$. Note that each pair of $\Im_i$ and L'_i is evaluated only once in two metrics.

Basic Settings. The corresponding datasets are divided into training part and testing part, which consist of 70% and 30% of the whole datasets selected randomly, each experiment is evaluated by 10 times and the average results are returned. The default values of $\alpha_{\vartheta}, \beta_{\vartheta}$ and γ_{ϑ} are all equal to $1/3$ at the beginning of model learning, λ_s and λ_t are set equal to 1. Meanwhile, τ_E is set to 200 metres, k is initialized as 20, 50, 80, and 100 separately. During the method practices, the multi-process programming is utilized to accelerate the whole evaluations while a total of 20 CPU cores are handled.

5.3 Experimental Results

Figure 4 and Fig. 5 have illustrated the precision and RMSE of varying k in two facility categories that are operated on Beijing and Shanghai respectively. It can be seen that the proposed MITLR significantly outperforms the other methods

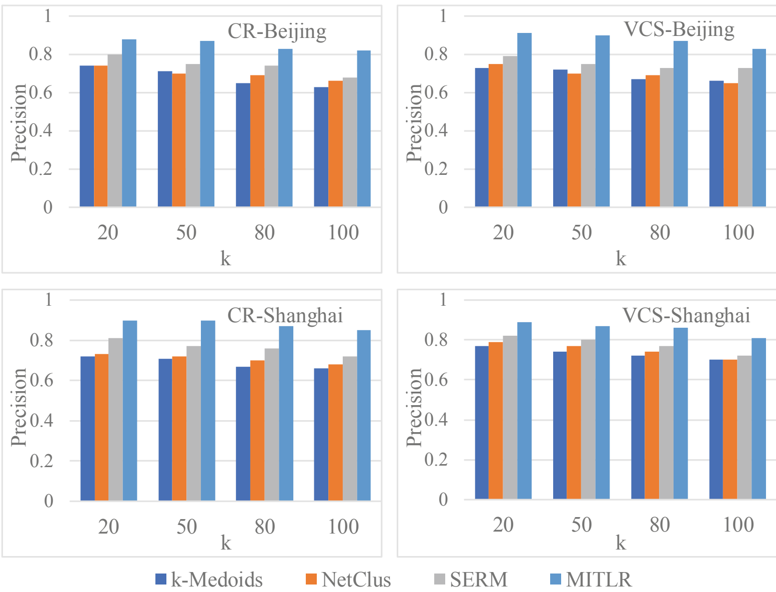

Fig. 4. Performance in terms of precision

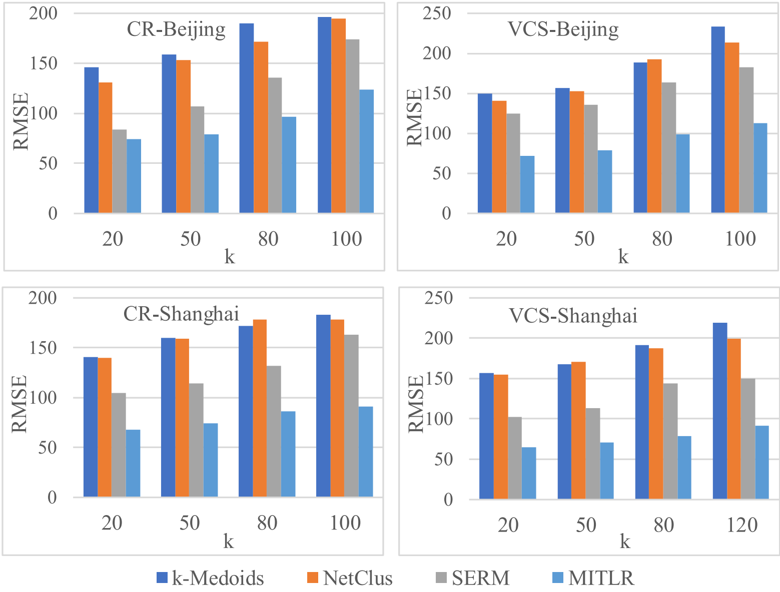

Fig. 5. Performance in terms of RMSE

under all circumstances in precision, the reasons would be analyzed in several important aspects, as we not only consider the effects of customer historical trajectories on facility placing, but also take advantage of the friend relationship based reciprocal recommendation, besides, the serious acknowledgements of service capacity improves the accuracies of prediction as well. Furthermore, the precision declines softly along with the raise of k, for the larger value of k, the hitting accuracy will experience diminishing returns.

Subsequently, we can see that our proposed model has better achievements compared with all competitors in RMSE, and the results has a reverse manner comparing with the precision, the reason could by analyzed from their definitions directly. In other words, if the value of precision is larger, the candidate locations recommended will be better represented by the corresponding facilities in road network, it also demonstrates that the larger RMSE will result in a worse performance on predicting contrarily.

We further investigate three feature weights in Fig. 6, when referring to α_ϑ, β_ϑ and γ_ϑ, the characteristic of social relevance is a principal factor on the evaluation of service utility especially in VCS, this is because that the majority of customers are more prone to be influenced by red envelopes or positive comments posted by their close friends when they are going to have consumptions at VCS, and vice versa.

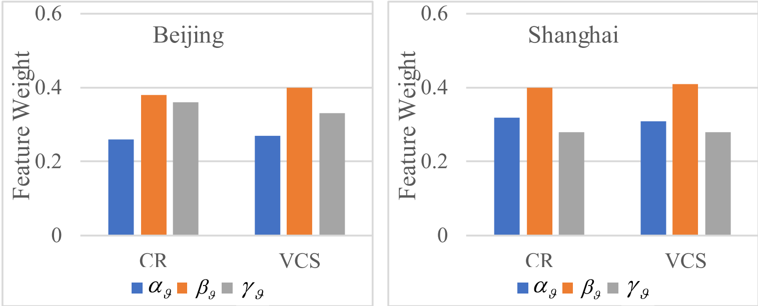

Fig. 6. Weights learned by MITLR

6 Conclusions

In this paper, we have defined a novel model MITLR for top-k facility location recommendation, it considers both spatio-temporal behaviors and social relationships of customers. In order to achieve effective query processing, CLTC-tree (forest) that combines candidate locations and customers are presented, and a query algorithm is also examined to obtain the results. Finally, extensive experiments with real datasets are performed to offer insights into the effectiveness.

References

1. Bao, J., He, T., Ruan, S., Li, Y., Zheng, Y.: Planning bike lanes based on sharing-bikes' trajectories. In: Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 1377–1386. ACM (2017)
2. Karamshuk, D., Noulas, A., Scellato, S., Nicosia, V., Mascolo, C.: Geo-spotting: mining online location-based services for optimal retail store placement. In: Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 793–801. ACM (2013)
3. Li, X., Čeikute, V., Jensen, C.S., Tan, K.L.: Trajectory based optimal segment computation in road network databases. In: Proceedings of the 21st ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems, pp. 396–399. ACM (2013)
4. Li, Y., Bao, J., Li, Y., Wu, Y., Gong, Z., Zheng, Y.: Mining the most influential k-location set from massive trajectories. In: Proceedings of the 24th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems, pp. 1–4. ACM (2016)
5. Mitra, S., Saraf, P., Bhattacharya, A.: Tips: mining top-k locations to minimize user-inconvenience for trajectory-aware services. IEEE Trans. Knowl. Data Eng. (2019)
6. Mitra, S., Saraf, P., Sharma, R., Bhattacharya, A., Ranuy, S., Bhandari, H.: NetClus: a scalable framework for locating top-k sites for placement of trajectory-aware services. In: 2017 IEEE 33rd International Conference on Data Engineering (ICDE), pp. 87–90. IEEE (2017)

7. Sun, Y., Huang, J., Chen, Y., Zhang, R., Du, X.: Location selection for utility maximization with capacity constraints. In: Proceedings of the 21st ACM International Conference on Information and Knowledge Management, pp. 2154–2158. ACM (2012)
8. Wei, L.Y., Zheng, Y., Peng, W.C.: Constructing popular routes from uncertain trajectories. In: Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 195–203. ACM (2012)
9. Yao, D., Zhang, C., Huang, J., Bi, J.: SERM: a recurrent model for next location prediction in semantic trajectories. In: Proceedings of the 2017 ACM on Conference on Information and Knowledge Management, pp. 2411–2414. ACM (2017)
10. Yin, H., Cui, B., Chen, L., Hu, Z., Zhang, C.: Modeling location-based user rating profiles for personalized recommendation. ACM Trans. Knowl. Discov. Data (TKDD) **9**(3), 1–41 (2015)
11. Yin, M., Li, Q., Liu, Y.: Study on herd behavior and its influence on consumer impulse buying behavior during online shopping festival: an empirical study based on social influence theory and herd behavior. China Bus. Mark. **33**(8), 99–107 (2019)