MC^{2}LS: Towards Efficient Collective Location Selection in Competition

(Extended Abstract)

Meng Wang[1], Mengfei Zhao[1], Hui Li[2,3,*], Jiangtao Cui[2], Bo Yang[1], Tao Xue[1]
[1]*School of Computer Science & The Shaanxi Key Laboratory of Clothing Intelligence, Xi'an Polytechnic University, China*
[2]*School of Computer Science and Technology, Xidian University, China*
[3]*Shanghai Yunxi Technology, China / *Corresponding author*
wangmeng@xpu.edu.cn, 961311016@qq.com, {hli, cuijt}@xidian.edu.cn, yangboo@stu.xjtu.edu.cn, xuetao@xpu.edu.cn

***Abstract*—Collective Location Selection (CLS) aims to identify *k* optimal sites for facility establishment to collectively maximize user attraction. Traditional CLS approaches often overlook user mobility and inter-facility competition, critical factors in real-world scenarios. This paper introduces MC^{2}LS, the first effort on CLS that addresses these gaps by considering user mobility and peer competition. Solving MC^{2}LS is non-trivial due to its NP-hardness. To overcome the challenge of pruning multi-point users with highly overlapping minimum boundary rectangles (MBRs), we develop a position count threshold and two square-based pruning rules. We propose IQuad-tree, a user-MBR-free index, to benefit the hierarchical and batch-wise properties of the pruning rules. We present an $(1-\frac{1}{e})$-approximate greedy solution to MC^{2}LS, and empirical studies demonstrate the superiority of our proposed solution over the state-of-the-art techniques.**

I. Introduction

The ubiquity of mobile internet and GPS devices have led to a surge in geo-tagged data, providing invaluable support for location decision-making. Large corporations and chains, often prioritize market share over individual facility impact, leading to the development of Collective Location Selection (CLS) strategies [1], which aims to identify a group of *k* optimal sites among candidates to collectively maximize user attraction. Traditional approaches measure location attractiveness based on spatial proximity, assuming users are static. However, a recent study on moving users [2], *k*-Collective Influential Facility Placement (*k*-CIFP), has challenged this single-point model through the use of multi-point semantics [3], where the influence relationship is defined as a cumulative probability of all positions of a user. Despite advancements, most CLS literature overlooks peer competitors due to the inherent complexity of modeling and evaluation, which can significantly impair the effectiveness of CLS models in competitive markets. To this end, we propose a novel and more practical CLS problem called **M**obility-oriented **C**ompetitive-based **CLS** (**MC^{2}LS**), which considers both mobility and competition factors.

II. Problem Definition

Following the mobility-aware influence criterion [2], a moving user o is represented by a set of r positions $o = \{p_1, p_2, \ldots, p_r\}$. As both candidate locations for opening new facilities (set $C = \{c_1, c_2, \ldots, c_n\}$) and existing facilities as competitors (set $F = \{f_1, f_2, \ldots, f_m\}$) can exert influence on users, we introduce a concept of ***abstract facility*** v, where $v \in C \cup F$. For user o, the ***cumulative influence probability*** is defined as $Pr_v(o) = 1 - \prod_{i=1}^{r}(1 - Pr_v(p_i))$, accounting for all positions of o. Here, $Pr_v(p_i) = PF(dist(v, p_i))$ denotes the independent probability that o is influenced by v at position $p_i \in o$, with PF being a distance-based probability function that can be adjusted to accommodate various influence patterns. In practice, businesses prioritize either high-quality or a broad range of users based on influence. A probabilistic threshold τ helps balance this trade-off by filtering desired users: If $Pr_v(o) \geq \tau$, v can influence user o. Given a user set $\Omega = \{o_1, o_2, \ldots, o_n\}$, the subset Ω_v consists of users influenced by v, and $inf(v) = |\Omega_v|$ quantifies v's ***influence*** value [3]. According to the *evenly split* competition model [4], facilities that attract the same user equally capture the influence. Considering competitors $F_o = \{f | Pr_f(o) \geq \tau \wedge f \in F\}$, the ***competitive influence of candidate*** c ***on*** o is calculated as $cinf(c, o) = \frac{1}{|F_o|+1}$. Then the ***competitive collective influence of*** G ($G \subseteq C$) is computed as $cinf(G) = \sum_{o \in \Omega_G} \frac{1}{|F_o|+1}$, where $\Omega_G = \{o | Pr_c(o) \geq \tau \wedge c \in G \wedge o \in \Omega\}$. Thus, **MC^{2}LS** aims to find an optimal candidate subset $G \subseteq C \wedge |G| = k$ to maximize its competitive collective influence.

III. IQuad-tree-based Solution to MC^{2}LS

We can prove MC^{2}LS to be isomorphic to the *Maximum k-Coverage* problem, whose NP-hard property results in exponential complexity when enumerating all combinations of k candidates out of n. Given that $cinf(c) = \sum_{o \in \Omega_c} cinf(c, o)$ enables the exact competitive influence of every candidate, greedy heuristics can find a near optimal solution in polynomial time. Even so, exhaustive calculation $cinf(c)$ for all candidates still accounts for prohibitive complexity $O((|C| + |F|)|\Omega|r)$. The challenge in designing efficient algorithms lies in the significant overlap of the activity regions (MBRs of positions) of moving users [3], making ***classical pruning rules fail to exclude irrelevant users***. Thus, the previous study [3] had to shift its focus to eliminating candidates rather than users. They employed an ***Influence Circle*** to measure multi-point-based influence relationships: an abstract facility v can (*resp.*, cannot) attract a user o if all of o's positions lie within (*resp.*, without) $\phi(v, mMR(\tau, r))$, which denotes a circle centered at v with radius $mMR(\tau, r) = PF^{-1}(1-(1-\tau)^{1/r})$.

2375-026X/25/$31.00 ©2025 IEEE
DOI 10.1109/ICDE65448.2025.00391

To tackle the challenge, we propose a user-MBR-free strategy via reverse deduction of $mMR(\tau, r)$, where $r = 1/\log_{1-\tau}(1 - PF(mMR(\tau, r)))$. This allows us a new position count threshold $\eta(\tau, PF, \hat{d}) = 1/\log_{1-\tau}(1 - PF(\hat{d}))$ to determine the influence relationship. If circle $\phi(v, \hat{d})$ encloses $\lceil \eta(\tau, PF, \hat{d}) \rceil$ positions of user o, abstract facility v must influence o. Based on $\eta(\tau, PF, \hat{d})$, we construct two square-based pruning rules to filter out users for a batch of candidates.

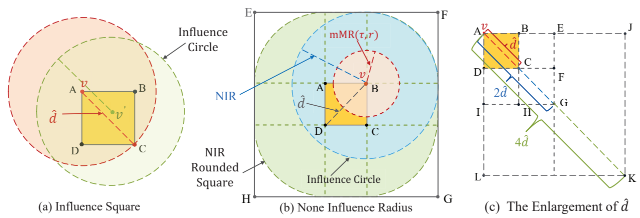

Fig. 1. Pruning rules and the square hierarchy index.

Influence Square (IS): As shown in Fig. 1(a), given a square $ABCD$ with diagonal lengths $\hat{d}$, any abstract facility v lying within the square can influence a user o, if at least $\lceil \eta(\tau, PF, \hat{d}) \rceil$ positions of o are covered by $ABCD$. This can be guaranteed by the aforementioned Influence Circle. **Non-Influence Radius (NIR)**: We define NIR as the maximum $mMR(\tau, r_{max})$ of all the users, where $r_{max} = \max\{r | r = |o| \wedge o \in \Omega\}$. In Fig. 1(b), we can draw a NIR rounded square based on ABCD. Then any abstract facility v located within square $ABCD$ cannot influence user o, if none of o's positions is inside the MBR (i.e., $EFGH$) of the NIR rounded square. Thus we can efficiently prune users who are necessarily influenced or not attracted in a given square.

Even if a user o fails to meet the IS rule, o may have additional positions outside the IS region, indicating that o could possibly satisfy an enlarged $\eta(\tau, PF, \hat{d})$ with longer $\hat{d}$. In Fig. 1(c), we extend $\hat{d}$ to $2\hat{d}$, $4\hat{d}$, and so on, and verify if the IS rule is satisfied. This manner motivates us to integrate the pruning rules into Quad-tree to develop an **IQuad-tree** (Influence Quad-tree) for indexing users and their positions.

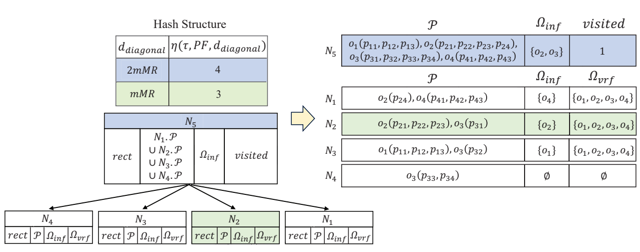

Fig. 2. The IQuad-tree index structure and an example.

The leaf node N^L of an IQuad-tree comprises entries in the form of $\langle rect, \mathcal{P}, \Omega_{inf}, \Omega_{vrf} \rangle$. $rect$ denotes the square of N^L with diagonal $\hat{d}$. In the *key-value* set $\mathcal{P}$, the *key* is user ID o_{id} and the *value* associates only the positions of o_{id} located inside $N^L.rect$. The sets Ω_{inf} and Ω_{vrf} hold users that are necessarily and possibly influenced by abstract facilities within $N^L.rect$. The users in $\Omega_{vrf} \setminus \Omega_{inf}$ need to be further verified.

For the entry $\langle rect, \mathcal{P}, \Omega_{inf}, visited \rangle$ of a non-leaf node N^N, $rect$ is the area enclosed by its four child squares. $N^N.\mathcal{P}$ is the union of $\mathcal{P}$ sets of child nodes. The set $N^N.\Omega_{inf}$ follows the enlargement shown in Fig. 1(c). The binary flag $visited$ indicates whether N^N has been traversed or not.

Based on IQuad-tree, we design a ***four-stage solution*** to MC^{2}LS. First, the IS and NIR pruning rules are utilized to batch-wise identify the influence relationships between users and abstract facilities in a node. The process is traversed recursively from leaf nodes to larger regions. Second, the validation $Pr_f(o) \geq \tau$ is performed for users not pruned yet. Third, we evaluate competitive influence of candidates $cinf(c)$. Finally, a greedy heuristic is used to gradually select k candidates with the maximum $cinf(c)$ at each step. As $cinf(\cdot)$ satisfies the condition of a ***submodular non-decreasing function***, the heuristic guarantees an $(1 - \frac{1}{e})$-approximation ratio.

IV. Summary of Experimental Evaluation

We have investigated the performance of our IQuad-Tree-based (IQT) solution compared to the adapted greedy version of k-CIFP [2], following the same settings over real datasets.

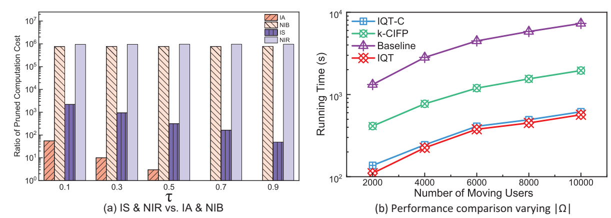

Fig. 3. Comparisons of performance and pruning rules.

As illustrated in Fig. 3(a), both the IS and NIR pruning rules work effectively. Compared to IS, which imposes relatively strict position counting within small squares, NIR exhibits superior performance owing to its larger pruning areas. Given a specific $\hat{d}$, as τ increases, $\eta(\tau, PF, \hat{d})$ grows while the NIR value declines, thereby reducing the efficiency of IS and enhancing that of NIR. In comparison to the candidate-pruning strategy [3], i.e., IA and NIB, our proposed IS and NIR rules are more effective due to their batch-wise property.

Fig. 3(b) demonstrates the scalability with varying $|\Omega|$, where IQT significantly reduces the running time by one order of magnitude compared to Baseline (exhaustively scanning all user-abstract facility pairs). The comparisons across varying parameters $|C|$, $|F|$, τ, k, $\hat{d}$ and r are qualitatively similar, where IQT exhibits the best efficiency, followed by IQT-C (without employing the IA rule), k-CIFP and Baseline.

References

[1] F. Chen, H. Lin, J. Qi, P. Li, and Y. Gao, "Collective-k optimal location selection," in *SSTD*, 2017, pp. 339–356.

[2] D. Li, H. Li, M. Wang, and J. Cui, "k-collective influential facility placement over moving object," in *MDM*, 2019, pp. 191–200.

[3] M. Wang, H. Li, J. Cui, K. Deng, S. S. Bhowmick, and Z. Dong, "Pinocchio: Probabilistic influence-based location selection over moving objects," *TKDE*, vol. 28, no. 11, pp. 3068–3082, 2016.

[4] R. Aboolian, O. Berman, and D. Krass, "Competitive facility location model with concave demand," *Eur. J. Oper. Res.*, vol. 181, no. 2, pp. 598–619, 2007.