Opening and Closing Dynamics of Competing Shop Groups over Spatial Networks

Takayasu Fushimi*, Kazumi Saito†‡, Kouzou Ohara§, Masahiro Kimura¶ and Hiroshi Motoda‖
*Tokyo University of Technology, Japan †Kanagawa University, Japan ‡RIKEN, Japan
§Aoyama Gakuin University, Japan ¶Ryukoku University, Japan ‖Osaka University, Japan

***Abstract*—We address the problem of opening and closing shops in group competitive environment, i.e., shops in the same group work cooperatively and those in different groups competitively, and analyze how the market share and location changes over time. We formulate a stochastic utility of each shop as a function of shop distance and attractiveness from which a market share is computed by weighting consumers buying power. We further place a constraint on the traveling time, which is crucial to reduce the computation time, and use a marginal gain of the market share as a measure to rank the candidate location. Using the real dataset of three convenience stores in four cities in Japan, we confirm that, despite the simplification we made in the model, rankings of the existing shops are shown to be high which implies that our model is reasonable. Further, comparison with the baseline gravity model shows that our model gives much more realistic results. Analyses of the dynamics of opening and closing shops indicate that the reasonable time-bound for walking is about 10 min., the market share of each group, thus total share, eventually increases although small, and the difference of the share within each group gradually becomes smaller, revealing that the spatial distribution of the shops in each group becomes more uniform.**

I. Introduction

With the advent of big data era, better use of business data has become crucially important to improve every aspect of economic activities. This is true in particular in domains where human activities are highly involved and its behaviors are complex and, thus, there does not exist a theoretical model that can predict the activities good enough. Typical examples are dynamics of opinion and community formations via complex interactions of people over time. A voter model and its variants are convenient tools to study these dynamics [1], [2]. This paper focuses on a different topic but in the same kind which is the dynamics of opening and closing of competing shop groups over a spatial network. There have been many studies that tackle the competitive facility [1] (shop) location problem and many models have been proposed so far. Some of them focus on the dynamics of the market share due to the change of the location and size of facilities [3], [4]. But, they do not analyze the dynamics of the market share over a certain period of time taking into account both opening and closing facilities in competitive environment.

Another line of studies related to this problem is about centrality measures. Especially, the closest to this work is group centrality [5], [6]. We introduced a closeness-based centrality measure to solve the facitlity allocation problem under the competitive environment in [6], but it only considered the problem of opening new shops and did not analyze the opening/closing dynamics. We have substantially extended that approach to reflect the stochastic nature of people's behavior. In this paper we report the opening and closing dynamics of competing shops.

To the best of our knowledge, this is the first attempt to study the dynamics of such business activities in group competitive environment under the context of spatial network analysis. Further, as is seen later, our framework is viewed as a new attempt toward integrating facility location problem studied in operational research and group centrality studied in network science, and can be extended to solve broader problems.

The shops we have in mind are those selling daily goods or service, e.g. convenience stores, fast food stores, drug stores, supermarkets, gas stations, etc., which are indispensable in daily life, and many shops with different groups are competing, e.g. 7-Eleven, Lawson, FamilyMart, MiniStop, etc. in case of convenience stores, McDonald, Burger KING, Wendy's, Five Guys, etc. in case of fast food stores. These shops are distributed throughout a city and located at places where the purchase demand of consumers is considered to be high.

Their locations must be easily accessible for the consumers to fulfill their demands as well as for the shops to maximize their profit. Those locations attract many shops. Shops in the same group work collaboratory and shops in different groups compete with each other. This brings us an interesting problem of where to open a new shop and where to close an old shop in group competitive environment. We propose a new model that can deal with this problem and seek to analyze its dynamics.

To solve this problem we first map the city we deal with to a road network, i.e., spatial network, and rank the nodes that are the candidates for opening or closing shops, based on consumers demands that are also allocated in nodes. We need to have a measure to evaluate each candidate node. There are many factors that affect consumers behavior. Among all, the following three are reported to be important, i.e., distance to shop, shop attractiveness and buying power [7]. Our measure can encompass these three factors, provided the respective data are available or learnable from data. To be more precise, we formulate a market share as a function of these factors and use its marginal gain/loss due to the opening/closing as a measure

[1] Facility, shop and store are used interchangeably in this paper.

IEEE/ACM ASONAM 2020, December 7-10, 2020
978-1-7281-1056-1/20/$31.00

to rank the candidate locations. The market share of a shop is computed by the weighted sum of the utility of shops in the same group. Shop utility is represented by a function of two key features: the traveling time and the shop attractiveness, which is similar to the function used in the gravity model [8]. The difference is that we use a softmax function as is often used in machine learning problem setting and the traveling time is constrained by a time-bound parameter. The market share thus defined naturally handles cooperative/competitive environment. Stochastic nature of the utility avoids all or nothing behavior, i.e., each consumer does not necessarily go to a specific shop, but instead, has a probabilistic choice to go among multiple shops. The final measure is the difference of the market share before and after a new shop is opened (marginal gain) or an old shop is closed (marginal loss). We rank the candidate nodes according to this measure and decide where to open a new shop (maximum gain) or to close an old one (minimum loss).

We use a real dataset of three convenience stores in four different cities in Japan for evaluation purpose and use a part of them as existing shops and estimate the ranking of the locations of the shops for which we have the opening record. We use the gravity model [8] as a baseline for comparison. We show from the simulation result that rankings computed by our model is much higher than those computed by the gravity model. This confirms that our model, which is very simple though, is more realistic and reasonable and that the time-bound constraint plays indeed an essential role to drastically improve the efficiency to compute the rank. We then analyze the dynamics of opening and closing shops in group competitive environment in which the total number of shops in each group is fixed. The major findings include that the reasonable time-bound is about 10 min., the market share of each group, thus total share, eventually increases although the amount is small, and the difference of the share within each group gradually becomes smaller, which reveals that the spatial distribution of the shops in each group converges to a more uniform distribution.

The paper is organized as follows. Section II describes related work. Section III gives our problem setting and explains our proposed method. In Section IV, we report and discuss experimental results using real-world data. Finally, Section V concludes this paper and addresses future work.

II. Related work

In this section, we briefly review existing work on the facility location problem and spatial network analysis using centrality measures.

A. Facility location problem

There are a vast amount of studies related to the facility location problem in which the goal is to determine optimal locations to build one or more new facilities, e.g., retail shops, in a continuous or discrete space. Use of a spatial network as in our work is viewed as a discrete space approach, but models developed for a continuous space can be easily applied to a discrete space. The gravity model [8] in which consumers choose one facility from possible candidates in a stochastic manner is a representative model that can apply to a continuous space and its utility function is similar to the one used in our proposed method. In the gravity model, the probability that a facility is chosen by consumers who live in a demand point is given by normalizing the raw utility of the facility over all possible ones, whereas our utility function gives the probability by using softmax function. The raw utility of a facility in the gravity model is proportional to its attractiveness such as the floor size, and inversely proportional to a power of the distance from the demand point to the facility, just as in the deterministic utility model [9]. The market share of the facility is computed as the weighted sum of the normalized utility values over all demand points using the buying power of each demand point as the weight. The major difference other than the shape of the function between the gravity model and our model is the time-bound constraint we adopted. The basic idea behind this constraint is in the same spirit of the cover-based approach [10] that assumes each facility has an area of influence with a certain radius and a consumer within the area chooses the facility. We demonstrate that the time-bound constraint we introduced can improve the resulting market share ranking, thus obtaining more reasonable results, and at the same time significantly reduce the computational cost.

Besides, using the traveling time instead of the physical distance as the travel cost in our work would allow us to handle more realistic situations. For example, we could flexibly mixture different transportation means such as walking and driving in our framework by introducing different time-bound constraints for different transportation means although we only consider the simplest case that people always walk to a facility in this work.

The gravity model is an extension of the proximity approach [11] that is the most primitive way to estimate the market share. In this approach, consumers choose the closest facility in physical distance. There exist various extensions other than the gravity model. Another extension is the cover-based approach.

These extensions were proposed independently and some of them were combined in a single model. The model we proposed in this work is simple enough, but can naturally incorporate all of them.

A similar problem has been addressed and discussed in the data base community, which is called the Reverse Nearest Neighbor (RNN) query problem [12], [13], [14]. This is a problem to find objects in the database whose most similar object is the given query, and can be classified into Monochromatic RNN and Bichromatic RNN. In particular, [12] deals with the problem of finding a query that maximizes the size of BRNN called MaxBRNN, and is said to be applicable to the problem of finding the location of a facility that attracts more customers. This is deterministic and can fall in the proximity approach described above.

As for the competitive setting in the facility location problem, Drezner et al. considered the interaction between competitors in chronological order and proposed a leader-follower model [4] in which facilities belong to either of two groups, leader or follower, and it is assumed the leader first builds new facilities or extends existing ones belonging to the own group and then the follower takes similar actions to maximize its profit according to the actions of the leader. The problem is to determine the best locations for the leader to build new facilities or extend existing ones so as to maximize its profit considering the follower will take the optimal actions. Unlike this work, our model considers both opening and closing facilities and can take into account multiple competing groups. The dynamics considering both opening and closing facilities over the planning horizon is considered by Current et al. in a non-competitive environment [3]. Zhang and Rushton also considered opening and closing one or more facilities in a competitive environment [15], but they do not analyze the dynamics of market share over a certain period of time.

B. Social media analysis

Another stream of studies on facility locations exploits posted data on social media including user demands and reputation information for shops and mobility data including location information of checked-in stores and WiFi accessed [16], [17], [18], [19]. Almost all of these studies combine or compare geographic and mobility features and predict optimal locations for retail stores to be opened through supervised learning techniques. Important geographic features used often are the density of existing stores, type of nearby located stores, competitiveness, real estate price, degree of traffic congestion, whereas important mobility features often used are the number of visited users and the ratio of people moving within or between areas. Noulas et al. performed predicting human mobility by using an extended gravity model as a link prediction problem over a place network where nodes are the popular places in a target city and links are direct transitions of users between them [20].

Unlike our study, these existing ones do not consider nor handle the competitiveness among groups and the dynamics of stores opening and closing mathematically in their models. It should be noted that our model can also exploit user mobility data obtained from social media and use them to learn more appropriate parameter values.

C. Centrality analysis of spatial networks

Centrality measures characterize the importance of nodes in networks, and are widely used for the analysis of spatial networks, e.g., road networks. For example, Crucitti et al. used the distribution of four kinds of centrality measures to capture structural similarity among certain areas in a road network [21]. Montis et al. revealed the characteristics of undirected networks whose nodes and weighted links correspond to local governments and the commuter traffic in between respectively by focusing on the relationship between the degree and the clustering coefficient [22]. Park et al. adopted three centrality measures, degree, closeness, and betweenness, and their entropy to analyze the differences in the topological structure of residential areas and downtown areas [23]. A centrality measure in principle quantifies the intrinsic nature that a node possesses for the entire network and its scores of neighboring nodes are likely to be similar to each other due to overlapping influences of neighboring nodes. Thus, nodes with high scores may not necessarily be suitable for the sites to allocate facilities because of the possible concentration and competition. To avoid this problem group centrality was introduced [5] in which the centrality score of a group of nodes is quantified. Our measure is considered to be a generalization of group centrality. We consider multiple groups that have cooperative intragroup relationship and the competitive inter-group relationship. To the best of our knowledge, this is the first study that accounts for such cooperative and competitive relationships to analyze a road network and applies it to the facility location problem. The model we proposed turns out very general, covers major approaches so far proposed separately and can be applied to a broader class of problems.

III. Proposed Models

In this section, we propose our model for predicting a market share from basic three factors described earlier, derive our measures to rank the candidate locations for opening and closing shops in group competitive environment, and discuss the computational complexity for obtaining these ranking values. We then propose our basic dynamic model of opening and closing shops.

A. Prediction model

Let $G = (\mathcal{V}, \mathcal{E})$ be a given simple undirected (or bidirectional) network without self-loops which is intended to represent a road network, where $\mathcal{V} = \{u, v, w, \cdots\}$ and $\mathcal{E} = \{e, \cdots\}$ are sets of nodes (junctions) and undirected links (roads between adjacent junctions), respectively. We also express each link e as a pair of nodes, *i.e.*, $e = (u, v)$. For each link $e = (u, v) \in \mathcal{E}$, we assign its traveling time $t(u, v)$ between these nodes. For each pair of nodes that does not have the direct connection, *i.e.*, $(u, w) \notin \mathcal{E}$, we define its traveling time $t(u, w)$ by the minimum traveling time over all possible paths between them. For each node $v \in \mathcal{V}$, we assume that v has some weight denoted by $b(v)$ which is intended to represent buying power (demand) by customers at node v in a road network. In our problem setting, we assume a fixed set of nodes $\mathcal{U} \subset \mathcal{V}$ which is intended to represent shops such as convenience stores on a road network. Also, we assume that each shop belongs to one of J groups, and shops in the same group work cooperatively, but those in different groups competitively. Hereafter, we express the group of each shop $u \in \mathcal{U}$ as an integer denoted by $\phi(u) \in \{1, \cdots, J\}$, and the set of shops of group j as $\mathcal{U}_j = \{u \in \mathcal{U} \mid \phi(u) = j\}$.

We assume that customers prefer to go to a shop with relatively small traveling times, as typically employed in the proximity approach described earlier. Moreover, we introduce a time-bound parameter (threshold) denoted by τ, like a cover-based approach, and define our bounded traveling time $t_b(u, v)$ between shop node $u \in \mathcal{U}$ and some other node $v \in \mathcal{V}$ as

$t_b(u,v) = t(u,v)$ if $t(u,v) \le \tau$; otherwise $t_b(u,v) = \infty$. We also introduce a function of shop attractiveness at node u denoted by $a(u)$ such as its floor area size. Then, by combining these factors with adequate parameters α and β as

$$q(u;v) = -\alpha \; t_b(u,v) + \beta \; a(u), \quad (1)$$

we can define the following stochastic utility value of a shop at node $u \in \mathcal{U}$ with respect to a customer at node $v \in \mathcal{V}$.

$$p(u;v,\mathcal{U}) = \frac{\exp(q(u;v))}{\sum_{w\in\mathcal{U}} \exp(q(w;v))}, \quad (2)$$

where recall that $t(u,v) > \tau$ implies $t_b(u,v) = \infty$ and $\exp(q(u;v)) = 0$. Note that by defining $\tilde{q}(u;v) = \log(a(u)/d(u,v)^{\lambda}) = -\lambda \log d(u,v) + \log a(u)$ without using the time-bound parameter τ, where $d(u,v)$ is the distance between u and v, and λ is a parameter that is determined based on the type of facility, we can reduce our utility function to those used in standard gravity-based approaches [8], which means that this stochastic utility value based on a softmax function naturally combines representative proximity, utility, cover-based, and gravity-based approaches. Thus, we can compute the following market share of a shop at node $u \in \mathcal{U}$ with respect to buying power.

$$f(u;\mathcal{U}) = \sum_{v\in\mathcal{V}} p(u;v,\mathcal{U}) \; b(v). \quad (3)$$

Then, we can newly propose and express the market share of group j as

$$f(\mathcal{U}_j;\mathcal{U}) = \sum_{u\in\mathcal{U}_j} f(u;\mathcal{U}).$$

In what follows, we propose a series of ranking measures based on group j's perspective.

B. Ranking measures

First, we consider the effect of closing an existing shop $x \in \mathcal{U}$ in group j's perspective. When closing the shop x, we can compute the following difference between utility values of a shop at node $u \in \mathcal{U}$ with respect to a customer at node $v \in \mathcal{V}$.

$$\begin{aligned} & p(u;v,\mathcal{U}\setminus\{x\}) - p(u;v,\mathcal{U}) \\ & \quad = \frac{\exp(q(u;v))}{\sum_{w\in\mathcal{U}\setminus\{x\}} \exp(q(w;v))} - \frac{\exp(q(u;v))}{\sum_{w\in\mathcal{U}} \exp(q(w;v))} \\ & \quad = \frac{\exp(q(x;v))}{\sum_{w\in\mathcal{U}} \exp(q(w;v))} \frac{\exp(q(u;v))}{\sum_{w\in\mathcal{U}\setminus\{x\}} \exp(q(w;v))} \\ & \quad = p(x;v,\mathcal{U}) p(u;v,\mathcal{U}\setminus\{x\}). \end{aligned}$$

Then, we can compute the loss of market share with respect to the group j,

$$\begin{aligned} g_j(x;\mathcal{U}) &= f(\mathcal{U}_j;\mathcal{U}) - f(\mathcal{U}_j\setminus\{x\};\mathcal{U}\setminus\{x\}) \quad (4) \\ &= \sum_{v\in\mathcal{V}} p(x;v,\mathcal{U}) \; b(v) \left(1 - \sum_{u\in\mathcal{U}_j\setminus\{x\}} p(u;v,\mathcal{U}\setminus\{x\})\right). \end{aligned}$$

Namely, we can regard Eq. (4) as a measure for closing old shops in group j's perspective.

Next, we consider the effect of opening a new shop $y \in \mathcal{V}$ in group j's perspective. When opening the facility y with shop attractiveness $a(y)$, we can compute the following difference between utility values of a shop at node $u \in \mathcal{U}$ with respect to a customer at node $v \in \mathcal{V}$.

$$\begin{aligned} & p(u;v,\mathcal{U}) - p(u;v,\mathcal{U}\cup\{y\}) \\ & \quad = \frac{\exp(q(u;v))}{\sum_{w\in\mathcal{U}} \exp(q(w;v))} - \frac{\exp(q(u;v))}{\sum_{w\in\mathcal{U}\cup\{y\}} \exp(q(w;v))} \\ & \quad = \frac{\exp(q(y;v))}{\sum_{w\in\mathcal{U}\cup\{y\}} \exp(q(w;v))} \frac{\exp(q(u;v))}{\sum_{w\in\mathcal{U}} \exp(q(w;v))} \\ & \quad = p(y;v,\mathcal{U}\cup\{y\}) p(u;v,\mathcal{U}). \end{aligned}$$

Then, we can compute the gain of market share with respect to the group j,

$$\begin{aligned} h_j(y;\mathcal{U}) &= f(\mathcal{U}_j\cup\{y\};\mathcal{U}\cup\{y\}) - f(\mathcal{U}_j;\mathcal{U}) \quad (5) \\ &= \sum_{v\in\mathcal{V}} p(y;v,\mathcal{U}\cup\{y\}) \; b(v) \left(1 - \sum_{u\in\mathcal{U}_j} p(u;v,\mathcal{U})\right). \end{aligned}$$

Namely, we can regard Eq. (5) as a measure for opening new shops in group j's perspective. We refer to our model that ranks shops by using Eqs. (4) and (5) as the measures respectively for closing an old shops and opening a new shops as the Time-bounded Group-perspective Stochastic (TGS) model.

Finally, we note that the following measures are to be used when no consideration is needed of both cooperation within the same groups and competition among the different groups in case of closing a shop and opening a shop, respectively. These are derived straightforwardly from Eqs. (4) and (5) above.

$$f(x;\mathcal{U}) = \sum_{v\in\mathcal{V}} p(y;v,\mathcal{U}) \; b(v) \quad (6)$$

$$f(y;\mathcal{U}\cup\{y\}) = \sum_{v\in\mathcal{V}} p(y;v,\mathcal{U}\cup\{y\}) \; b(v) \quad (7)$$

Hereafter, this ranking measure is referred to as the NGP (No Group Perspective) method, where a ranking measure based on a standard gravity model can be obtained by simply replacing $q(u;v)$ of the NGP method to $\tilde{q}(u;v)$. We also note that our proposed ranking measure can be reduced to a node selection criterion for a standard group closeness centrality by setting $\alpha = \infty$, $\beta = 0$, $b(v) = 1$, and $J = 1$, where note that $\alpha = \infty$ means the nearest neighbour assignment. This implies that our proposed model can be regarded as a natural extension of notion of group centrality.

C. Computational complexity

We discuss the computational complexity of computing the above measures under the situation that the number of groups J is substantially smaller than the numbers of nodes and shops, $N = |\mathcal{V}|$ and $M = |\mathcal{U}|$. To this end, for a given time-bound τ, we define the set of neighbour nodes of $u \in \mathcal{U}$ as $\mathcal{V}(u,\tau) =$

$\{v \in \mathcal{V} \mid t(u,v) < \tau\}$, and define the average number of neighbour nodes as

$$L(\tau) = \frac{\sum_{u\in\mathcal{U}} |\mathcal{V}(u,\tau)|}{|\mathcal{U}|} \approx \frac{\sum_{v\in\mathcal{V}} |\mathcal{V}(v,\tau)|}{|\mathcal{V}|}.$$

Here, we should note that the scopes of the first summation appearing in the right-hand-side of Eqs. (4) and (5) can be replaced from $\mathcal{V}$ to $\mathcal{V}(x,\tau)$ and $\mathcal{V}(y,\tau)$, respectively.

In order to compute the stochastic utility value defined in Eq. (2) for every shop node $u \in \mathcal{U}$, we need to compute $\mathcal{V}(u,\tau)$ by performing the best-first search starting from each u with the computational complexity of $O(M \times L(\tau) \times \log L(\tau))$. Then, we can compute all of the stochastic utility values with the computational complexity of $O(M \times L(\tau))$, together with

$$R_j(v) = \sum_{w\in\mathcal{U}_j} \exp(q(w;v))$$

for every j and

$$R(v) = \sum_{w\in\mathcal{U}} \exp(q(w;v)).$$

Also, by noting that

$$\sum_{u\in\mathcal{U}_j\setminus\{x\}} p(u;v,\mathcal{U}\setminus\{x\}) = \frac{R_j(v) - \exp(q(x;v))}{R(v) - \exp(q(x;v))},$$

we can compute the loss of the market share defined in Eq. (4) for every shop node $x \in \mathcal{U}$ with the computational complexity of $O(M \times L(\tau))$. Thus, the total computational complexity of obtaining our shop closing measure becomes $O(M \times L(\tau) \times \log L(\tau))$, which is independent of N in virtue of time-bound τ.

In order to compute our shop opening measure defined in Eq. (5) for every shop candidate node $v \in \mathcal{V}$, we need to additionally compute $\mathcal{V}(v,\tau)$ by performing the best-first search starting from each v with the computational complexity of $O(N \times L(\tau) \times \log L(\tau))$. Then, by noting that

$$p(y;v,\mathcal{U}\cup\{y\}) = \frac{\exp(q(y;v))}{R(v) + \exp(q(y;v))}$$

and

$$\sum_{u\in\mathcal{U}_j} p(u;v,\mathcal{U}) = \frac{R_j(v)}{R(v)},$$

we can compute the gain of the market share defined in Eq. (5) for every shop node $y \in \mathcal{V}$ with the computational complexity of $O(N \times L(\tau))$. Thus, the total computational complexity of obtaining our opening shop measure becomes $O(N \times L(\tau) \times \log L(\tau))$. Therefore, the total computational complexity of obtaining these measures simultaneously becomes $O((M + N) \times L(\tau) \times \log L(\tau))$.

TABLE I
NUMBER OF NODES (SHOPS OR JUNCTIONS) IN EACH CITY.

City	g1	g2	g3	No Shops	Total
Hachioji	95	76	36	12,117	12,324
Sagamihara	132	77	79	18,822	19,110
Shizuoka	118	106	65	30,752	31,041
Yokohama	492	417	428	123,379	124,716

D. Dynamic model

We assume the following basic model for this analysis. We first assign an initial set of shops $\mathcal{U}^{(t)}$ at time $t = 0$. We then consider repeating the following steps based on the measures defined in Eqs. (4) and (5) without changing the number of total shops $|\mathcal{U}^{(t)}|$. After randomly selecting one of J groups as j, we compute a closing shop as $\hat{x} = \arg\min_{x\in\mathcal{U}_j^{(t)}}\{g_j(x;\mathcal{U}^{(t)})\}$, and an opening shop as $\hat{y} = \arg\max_{y\in\mathcal{V}\setminus\mathcal{U}^{(t)}}\{h_j(y;\mathcal{U}^{(t)})\}$, respectively, and then update the locations of these shops as $\mathcal{U}^{(t+1)} \leftarrow \mathcal{U}^{(t)} \setminus \{\hat{x}\} \cup \{\hat{y}\}$ together with setting $t \leftarrow t+1$. Here we assume that the shop attractiveness is inherited as $a(y) \leftarrow a(x)$.

As a basic property of our proposed model, especially in case that the shop attractiveness values are the same, we conjecture that the market shares of individual shops eventually will likely to be distributed uniformly because the model opens a new shop at location of the maximum market share after closing the shop with minimum market share in group j's perspective. Evidently, we can extend our model to a wide variety of directions. For instance, the shop closing/opening operation with group j might happen proportionally to the market share of group j. However, since exploring various situations is beyond the scope of this paper, we evaluate the basic property in our experiments.

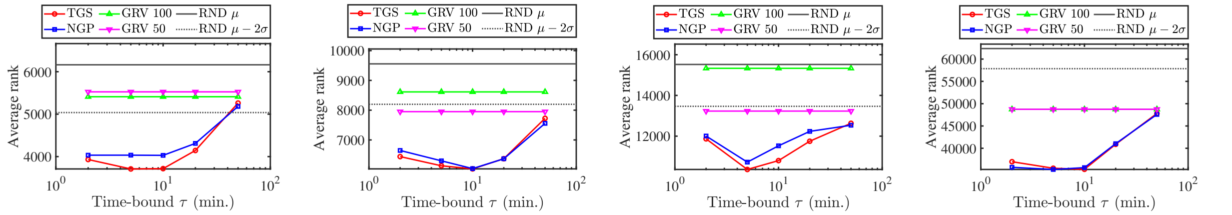
IV. EXPERIMENTS

A. Dataset and experimental settings

In our experimental evaluation, four cities, Hachioji, Sagamihara, Shizuoka, and Yokohama, are the target areas. We collected the road structure in these areas from OpenStreetMap (OSM) [2], and extracted all junctions and roads of each city. We then constructed a spatial network with the junctions as the nodes and the roads between the junctions as the links by following a standard formulation of road networks, such as those presented by SNAP (Stanford Network Analysis Project) [3]. We also collected the location information of actual stores of the three major convenience store chains in Japan (g1:7-Eleven, g2:FamilyMart, g3:Lawson) from NAVITIME [4]. Table I shows the number of junction nodes where convenience stores of the three groups are located and the number of normal junction nodes where no shops are located.

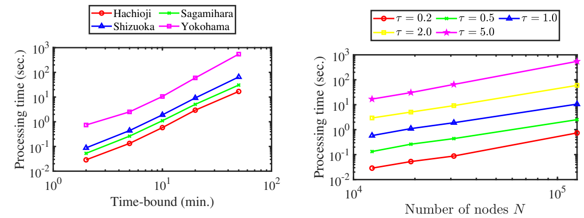
In our experiments, we set the buying power $b(v)$ to the daytime population, i.e., the number of nearby employees,

[2]https://www.openstreetmap.org/
[3]http://snap.stanford.edu/data/index.html
[4]https://www.navitime.co.jp/category/



(a) Hachioji (b) Sagamihara (c) Shizuoka (d) Yokohama

Fig. 1. Prediction accuracy.



(a) w.r.t time-bound τ (b) w.r.t #nodes N

Fig. 2. Processing time.

obtained from the 2015 census population aggregation data [5]. We assumed that a customer moves at 100 meter per minute on foot, and set the time-bound parameter τ to 2, 5, 10, 20, and 50 minutes as reasonable candidates. In this experiment we assumed that people walk to the shops for simplicity, but in reality some may drive. To cope with this, different τ can be used for different areas. Here we set parameter α to a relatively small value, i.e., $\alpha = 0.1$, so that stochastic utility value within the scope of the time-bound parameter does not become extremely small for $\tau = 50$ minutes which is considered maximum for walking. On the other hands, in our experiments using the convenience stores, we assumed that the shop attractiveness values are the same because their floor area sizes and the other properties are quite similar for all the shops. Then, we can cancel out the term $\beta a(u)$ from our stochastic utility function defined in Eq. (2) regardless of β value.

B. Compared methods

In our experiments, we evaluate our model TGS against the performance of the following three methods NGP, GRV and RND in terms of prediction accuracy.

- NGP is a special case of our model that does not consider both cooperative and competitive relationships as defined in Eq. (7).
- GRV is the gravity model which we consider is the state-of-the-art baseline and indeed is a reduced case of our model. It does not consider the time constraint and the form of utility function is different. However, straightforwardly calculating the utility function of the gravity model takes a huge amount of computation time even for networks we used in our experiments. Thus, we introduce our time-bound mechanism to the gravity model and set a large value for τ.
- RND is a random method, where all nodes in $\mathcal{V}$ are randomly sorted.

C. Prediction accuracy

First, we evaluate our model in terms of prediction accuracy. Since we know the opening dates for 40 shops in Hachioji, 66 in Sagamihara, 76 in Shizuoka, and 255 in Yokohama, we evaluate our model based on the average rank of these stores in the ranking of the gain of market share that could be brought when they are assumed to be newly opened [6]. More specifically, let $\mathcal{U}$ be a set of shops in a city, $\mathcal{U}^+ \subset \mathcal{U}$ be a set of shops with opening date available, and $m = |\mathcal{U}^+|$ be its size. Then, the average rank for a given method ψ is calculated as follows:

$$avg\_rank_\psi = \frac{1}{m} \sum_{y \in \mathcal{U}^+} r_\psi(y; \mathcal{U}), \tag{8}$$

where $r_\psi(y; \mathcal{U})$ stands for the rank of $y \in \mathcal{U}^+$ over $\mathcal{V}$ in terms of the gain of market share calculated by the method ψ assuming that each shop in $\mathcal{U}^+$ was newly opened. For our proposed method TGS, $r_{TGS}(y; \mathcal{U})$ is given by the rank of y over resulting values of $h_{\phi(y)}(y'; \mathcal{U}(y))$ for each $y' \in \mathcal{V} \backslash \mathcal{U}(y)$, where $\mathcal{U}(y)$ consists of shops that have been opened before y is opened.

Figure 1 shows the average ranks of our model and three other models, with respect to the time-bound parameter $\tau \in \{2, 5, 10, 20, 50\}$ described above. In Fig. 1, the red and blue lines show the average ranks of our proposed model TGS that considers both cooperation within the same groups and competitions among the different groups as defined in Eq. (5), and the compared model NGP that does not consider both relationships as defined in Eq. (7), respectively. The green and magenta lines are those of gravity model GRV, where we set the time-bound parameter as $\tau = 100$ and $\tau = 50$, respectively. As for the parameter λ of the traveling cost $d(u, v)^\lambda$ of the utility function of the gravity model, we set $\lambda = 3$ according to the known results [4]. The black line is the random model RND for which the average rank of nodes in $\mathcal{U}^+$ is $\mu = \frac{N+1}{2}$, where $N = |\mathcal{V}|$. Its standard deviation is $\sigma = \sqrt{\frac{N^2-1}{12m}}$. The dotted black line indicates $\mu - 2\sigma$ of RND.

[5] https://www.e-stat.go.jp/gis/

[6] Here we should note that the closing date information was not available in the above datasets.

From Fig. 1, we confirm that the average rank of TGS is the highest (smallest value) for most of the different time-bound values in all cities. Many cities have the highest average rank at near the time-bound $\tau = 10$. It should be noted that the difference of the average rank of TGS from that of RND is statistically significant, i.e., the rank is more than 2σ higher for all the cities. Here we should emphasize that the good performance around the time-bound $\tau = 10$ (1km in distance) coincides with our intuition. We can further see that the proposed model has slightly better performance than NGP, confirming the merit of explicitly taking in both cooperative and competitive factors. Moreover, in case that when the time-bound is small or large, i.e., $\tau = 2$ or 50, we can see that the performance of both TGS and NGP becomes comparable, which implies that both relationships do not work adequately because the constraint is too tight and too loose, respectively. It is evident that the average rank of TGS is substantially higher than those of GRV 100 and GRV 50 when τ is small, e.g., less than 30. This is attributed to the utility function of the gravity model that has no constraint on distance and thus consider the utility value even for a node that is far away from a shop. Setting the parameter λ of $\frac{1}{d(u,v)^{\lambda}}$ to a large value can reduce the effect of such distant nodes in the resulting utility value, but it reduces the utility of near nodes, too. Thus, adjusting the value of λ cannot replace the time-bound constraint, and it is crucial to explicitly impose a constraint on the traveling cost in order to obtain reasonable results.

In summary, we can say from these results that choosing an appropriate time-bound, say 10 in this case, allows us to improve the performance of the model and obtain a more reasonable result from the group perspective.

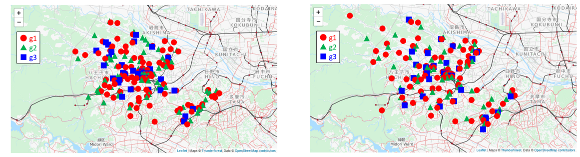
D. Computational efficiency

Next we evaluate our model in terms of computational efficiency. Figure 2 shows the processing time (sec.) of computing our measures Eqs. (4) for closing and (5) for opening in this order with respect to the time-bound parameter τ (Fig. 2(a)) and the number of nodes N (Fig. 2(b)) on log-log scale. Complexity analysis in Section III-C shows that the computation time is $O((M+N) \times L(\tau) \times \log L(\tau))$. We can assume that $L(\tau)$ increases quadratically with τ because road network extends planarly.

From Fig. 2(a), we can see that, as τ increases, regardless of cities, the computation time increases almost linearly with a slope around 2, which supports that $L(\tau)$ is a quadratic function, together with the validity of our analyses.

We note that computation time for the reasonable τ (about 10) is less than 5 seconds, which is very efficient. From Fig. 2(b), we can see that the computation time increases almost linearly with a slope around 1, which also supports the validity of our analyses under the setting of $N \gg M$.

From these results, it can be concluded that the time-bound constraint has remarkable effect on reducing the computation time if we have to rank all the nodes which is the case here. Remember that we cannot obtain the results of the gravity model within a reasonable time without using the time-bound constraint.



(a) Initial state $\mathcal{U}^{(0)}$. (b) After 100 cycles $\mathcal{U}^{(100)}$.

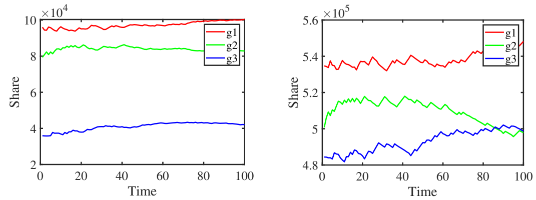
Fig. 3. Spatial distibution of convenience stores in Hachioji.

Finally we evaluate how the shop location changes over time as old shops are closed and new shops are opened as described in Section III-D. We repeated closing/opening cycle 100 times after setting an initial state $\mathcal{U}^{(0)}$ to the current shop locations, Figure 3 is the result of Hachioji, where Figs 3(a) and (b) show the locations for $\mathcal{U}^{(0)}$ and $\mathcal{U}^{(100)}$, respectively. It is clear that the shop is distributed more uniformly after 100 cycles and this matches our conjecture. Note that 7-Eleven (g1) which has the largest number of shops extends their shops to outskirts in the city near mountains. Other shops which has less number of shops are not affordable to do so and focus more in the city center. This also support that our modeling is reasonable.

Figure 4 is the changes in the market share of each group $f(\mathcal{U}_j^{(t)}; \mathcal{U}^{(t)})$ over time t for Hachioji and Yokohama. Other two cities exhibits similar changes. We see that there are some transient behaviors during which some share increases while some other decreases. In the end we can say that each group's share converges to a slightly larger value than its respective initial because each shop tries to minimize the loss and maximize the gain at each cycle. Figure 5 show how the difference of the market shares within each group $\max_{u \in \mathcal{U}_j^{(t)}} f(u; \mathcal{U}^{(t)}) - \min_{u \in \mathcal{U}_j^{(t)}} f(u; \mathcal{U}^{(t)})$ changes over time t. It is clear that as time goes, the gain decreases because the best location has been taken already and each group has to choose the next best in the remaining locations.

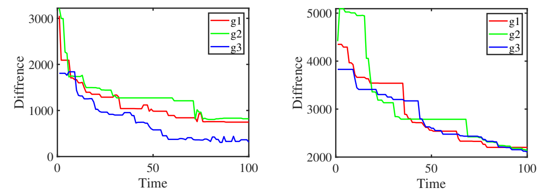
V. Conclusion

In this paper we addressed the problem of opening and closing shops in competitive environment and analyzed how the market share and location changes over time in a spatial network. We proposed to use a market share which accounts for major factors affecting consumers' behavior: distance to shop, shop attractiveness and buying power, and its marginal gain/loss due to the opening/closing as a measure to rank the candidate locations. The simulation results revealed that the rankings of the later opened shops computed by our model is much higher than those computed by the baseline gravity model. We found that the time-bound is crucial to obtain the realistic results as well as to gain substantial computational efficiency. This confirmed that, despite the simplification we made in the model, our model can beat the state-of-the-art. We then analyzed the dynamics of opening and closing shops in group competitive environment. The major findings include that the reasonable time-bound is about 10 min. for which



(a) Hachioji (b) Yokohama

Fig. 4. Changes in market share by group.



(a) Hachioji (b) Yokohama

Fig. 5. Changes in market share differences within a group.

the prediction accuracy is best, the market share of each group, thus total share, undergoes through transient states but eventually increases although its amount is small, and the difference of the share within each group gradually becomes smaller, revealing that the spatial distribution of the shops in each group becomes more uniform. We believe that this is the first attempt that integrates facility location problem and group centrality both of which have so far been studied in isolation, the former in operation research and the latter in network science. Our framework is general and is not restricted to facility location problem. We plan to extend this approach to more broader problems, e.g. information diffusion, opinion formation, etc.

ACKNOWLEDGMENTS

This material is based upon work supported by JSPS Grant-in-Aid for Scientific Research (C) (No.20K11940) and Early-Career Scientists (No.19K20417).

REFERENCES

[1] E. Even-Dar and A. Shapira, "A Note on Maximizing the Spread of Influence in Social Networks," in *Internet and Network Economics*, ser. Lecture Notes in Computer Science, X. Deng and F. Graham, Eds. Springer Berlin / Heidelberg, 2007, vol. 4858, pp. 281–286.

[2] M. Kimura, K. Saito, K. Ohara, and H. Motoda, "Learning to predict opinion share and detect anti-majority opinionists in social networks," *Journal of Intelligent Information Systems*, vol. 41, pp. 5–37, 08 2013.

[3] J. Current, S. Ratick, and C. ReVelle, "Dynamic facility location when the total number of facilities is uncertain: A decision analysis approach," *European Journal of Operational Research*, vol. 110, no. 3, pp. 597 – 609, 1998.

[4] T. Drezner, Z. Drezner, and P. Kalczynski, "A leader–follower model for discrete competitive facility location," *Computers & Operations Research*, vol. 64, pp. 51 – 59, 2015.

[5] M. G. Everett and S. P. Borgatti, "The centrality of groups and classes," *Journal of Mathematics Sociology*, vol. 23, no. 3, pp. 181–201, 1999.

[6] T. Fushimi, S. Okubo, and K. Saito, "Facility location problem on network based on group centrality measure considering cooperation and competition," in *Proceedings of the 8th International Conference on Complex Networks and Their Applications*, 2019, pp. 64–76.

[7] T. Drezner, "A review of competitive facility location in the plane," *Logistics Research*, vol. 7, 2014.

[8] ——, "Optimal continuous location of a retail facility, facility attractiveness, and market share: An interactive model," *Journal of Retailing*, vol. 70, no. 1, pp. 49–64, 1994.

[9] ——, "Locating a single new facility among existing unequally attractive facilities," *Journal of Regional Science*, vol. 34, no. 2, pp. 237–252, 1994.

[10] T. Drezner, Z. Drezner, and P. Kalczynski, "A cover-based competitive location model," *The Journal of the Operational Research Society*, vol. 62, no. 1, pp. 100–113, 2011.

[11] H. Hotelling, "Stability in competition," *The Economic Journal*, vol. 39, no. 153, pp. 41–57, 1929.

[12] R. C.-W. Wong, M. T. Özsu, P. S. Yu, A. W.-C. Fu, and L. Liu, "Efficient method for maximizing bichromatic reverse nearest neighbor," *Proc. VLDB Endow.*, vol. 2, no. 1, pp. 1126–1137, Aug. 2009. [Online]. Available: https://doi.org/10.14778/1687627.1687754

[13] Y. Gao, B. Zheng, G. Chen, W. Lee, K. C. K. Lee, and Q. Li, "Visible reverse k-nearest neighbor queries," in *2009 IEEE 25th International Conference on Data Engineering*, 2009, pp. 1203–1206.

[14] Y. Gao, J. Yang, G. Chen, B. Zheng, and C. Chen, "On efficient obstructed reverse nearest neighbor query processing," in *Proceedings of the 19th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, ser. GIS'11. New York, NY, USA: Association for Computing Machinery, 2011, pp. 191–200. [Online]. Available: https://doi.org/10.1145/2093973.2094000

[15] L. Zhang and G. Rushton, "Optimizing the size and locations of facilities in competitive multi-site service systems," *Computers & Operations Research*, vol. 35, no. 2, pp. 327 – 338, 2008.

[16] M. Xu, T. Wang, Z. Wu, J. Zhou, J. Li, and H. Wu, "Demand driven store site selection via multiple spatial-temporal data," in *Proceedings of the 24th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, ser. SIGSPATIAL'16. New York, NY, USA: Association for Computing Machinery, 2016. [Online]. Available: https://doi.org/10.1145/2996913.2996996

[17] F. Wang, L. Chen, and W. Pan, "Where to place your next restaurant? optimal restaurant placement via leveraging user-generated reviews," in *Proceedings of the 25th ACM International on Conference on Information and Knowledge Management*, ser. CIKM'16. New York, NY, USA: Association for Computing Machinery, 2016, pp. 2371–2376. [Online]. Available: https://doi.org/10.1145/2983323.2983696

[18] J. Lin, R. Oentaryo, E.-P. Lim, C. Vu, A. Vu, and A. Kwee, "Where is the goldmine? finding promising business locations through facebook data analytics," in *Proceedings of the 27th ACM Conference on Hypertext and Social Media*, ser. HT'16. New York, NY, USA: Association for Computing Machinery, 2016, pp. 93–102. [Online]. Available: https://doi.org/10.1145/2914586.2914588

[19] J. Lian, F. Zhang, X. Xie, and G. Sun, "Restaurant survival analysis with heterogeneous information," in *Proceedings of the 26th International Conference on World Wide Web Companion*, ser. WWW'17 Companion. Republic and Canton of Geneva, CHE: International World Wide Web Conferences Steering Committee, 2017, pp. 993–1002. [Online]. Available: https://doi.org/10.1145/3041021.3055130

[20] A. Noulas, B. Shaw, R. Lambiotte, and C. Mascolo, "Topological properties and temporal dynamics of place networks in urban environments," in *Proceedings of the 24th International Conference on World Wide Web*, ser. WWW'15 Companion. New York, NY, USA: Association for Computing Machinery, 2015, pp. 431–441. [Online]. Available: https://doi.org/10.1145/2740908.2745402

[21] P. Crucitti, V. Latora, and S. Porta, "Centrality Measures in Spatial Networks of Urban Streets," *Physical Review E*, vol. 73, no. 3, pp. 036 125+, Mar. 2006.

[22] D. A. Montis, M. Barthelemy, A. Chessa, and A. Vespignani, "The Structure of Interurban Traffic: A Weighted Network Analysis," *Environment and Planning B: Planning and Design*, vol. 34, no. 5, pp. 905–924, 2007.

[23] K. Park and A. Yilmaz, "A Social Network Analysis Approach to Analyze Road Networks," in *Proceedings of the ASPRS Annual Conference 2010*, 2010.