Measuring the Intensity of Local Units' Locational Concentration with Regard to the Neighborhood Externality with GIS

Hiromi Mori[1,4], Noriaki Sakamoto[2], Hirokazu Hasegawa[3]
[1]Hosei University, Tokyo, JAPAN
[2]Hosei University, Tokyo, JAPAN
[3]Niigata Municipal Office, Iiigata city, JAPAN
Corresponding author: Hiromi Mori, e-mail: hiromim@hosei.ac.jp

Abstract

The intensity of concentration of enterprises and local units used to have been measured by the density, like population density, defined by their number over a designated unit of area. We will propose a new measure which can account also for the concentration status of neighboring areas.

We employed for this study the telephone directory database: "Yellow pages" provided by the NTT Business Information Service, Inc. We chose the Hachioji city (in Tokyo metropolitan area) as a survey field and January 2012 as reference date.

Firstly, we identified the local units which are supposed to be in live status by carrying out some adjustments required for using this directory data and with due matching procedures. Secondly, we obtained coordinate information for respective local units with address matching. Since several local units often accommodate one and the same address in urban areas, these geographical points give manifold scores in terms of the number of local units which share the same address. By using the buffering method, we then re-scored the spots which will multiply the intensity of the spots by taking into account the concentration intensity of the neighboring areas that fall within the designated scale of buffer. Finally, we mapped the areas by descending manner according to the calculated scores which are expected to delineate zones in terms of local units' concentration by also taking into account the intensity of concentration in the neighboring subareas.

Keywords: buffering, externality , GIS, local unit, locational concentration

1. Introduction

Sites where local units deploy their business activities have cardinal importance for their successful performance. Businesses opt to decide sites for new shops, offices, or factories by assessing negative external effects brought about by neighboring competitors while expecting positive external effects generated by locational concentration of wider spectrum of business activities.

Analyses based on official statistics used to have employed density to assess the intensity of local units' concentration. Since density gives a leveled measure in terms of their concentration in a specified area, ratios for smaller areas tends to lack stability and the resolving power diminishes as area in concern expands.

When geographical coordinates (longitude and latitude) are applicable, one can assess the manner how local units concentrate with the help of spatial aggregation of the data. We propose in this paper the two measures to assess the magnitude of local units' concentration of respective sites where they locate by taking external effects caused by surrounding subareas into consideration. By regarding coordinates as carrier for local units we first give score for respective locational points in terms of

their concentration, second re-score the measure by taking conditions in neighboring areas into consideration, and finally portray on map how highly scored business areas distribute in the city by degree of concentration.

2. Data and target area

We used "yellow page" telephone directory database data provided by a private data provider to identify the location of local units. Yellow pages do not necessarily carry telephone number of whole existing local units. The database records only local units which requested to put ones names on the list among wired phones and toll-free numbers. Thus it does not cover telephones such as mobiles and IP phones, while local units with several numbers also appear in the database. Despite these shortcomings, the yellow page data enjoy smart coverage of local units because of advertising effect. Furthermore it also gives local units' names, addresses and types of businesses which meet manifold analyses.

The area in concern in this study is Hachioji : a western suburban city of Tokyo with about 550,000 population. The city has a wider spectrum of industries.

3. Preprocessing the data

As yellow page database contains some cases which have already expired or forward new numbers due to the local units' outside move, we cleaned the telephone number before processing the data. Hereafter N_orig and N_clnd denote original and cleaned datasets, respectively. Cases such as local units' moving outs, expiries, channel errors of the phones are deleted from N_orig to generate N_clnd by cleaning the data.

4. Identification of active cases

We employed statistical data matching to identify possible candidates of active cases by using N_clnd as of 2011.1 (N_clnd(11)) and those of 2012.1 (N_clnd(12)). Three variables (numbers, names of firms and addresses) were used in the matching process.

The matching produces 8 different matching patterns: triple match (A), double match (B,C,D), single match (E,F,G) and the unmatched. Among cases which fall in the final category there exists the unmatch (H) with N_clnd(12) as recipients and (I) with N_clnd(11) as recipients. Category (I) possibly involve local units which have closed or moved out during the reference period. Among single matches as for the number (F) or address (G) match it is possible that N_clnd(11) may involve the closed and moved out cases. Moreover, among local units which fall in these categories there may exist also cases caused by the withdrawals on request or by using a new phone numbers which are not covered by yellow pages. In the last two categories of cases local units do exist despite the absence of relevant records in N_clnd(12).

In order to identify absence or presence of the local units we conducted a special mail survey for the 1,144 unidentified units which are in the categories (I), (F) and (G). According to the survey results, 81 local units turned out to be active. We finally regarded the number of active local units as 14,961 as for 2012.1.

5. Analyses

(1) obtaining the coordinate information

We obtained coordinates for 14,961 local units by means of address matching. In cases when several local units share the same address e.g. by renting offices in multitenant buildings, yellow page database gives the same addresses for units which share the building. These local units have consequently the identical coordinates.

(2) assessing the concentration intensity by accounting the externality

(a) neighboring locational concentration (category 1)

We first counted the cases of identical coordinates based on the data given by address matching for respective active local units. The counted scores are identical with the results obtained also by spatial aggregation with no buffer. In cases when several local units share the same address in the yellow pages, they locate closely each other often sharing the same building. The score may indicate the intensity of local units' concentration at the respective spots where they exist which can represent a kind of "vertical concentration". We term it in this paper "the neighboring location concentration of category 1" (NLCC1).

It may be worth noting here that the NLCC1 does not necessary mean the potential capacity of each facility to accommodate tenants but does represent the concentration attributes given by the number of active local units. The score is defined simply by the number of local units carried by respective locational points and thus it in its nature differs from density standardized in terms of the size of areas.

(b) neighboring locational concentration (category 2)

We then re-scored the NLCC1 by appropriate buffering from respective coordinates. In cases when neighboring locational points fall within the designated size of buffers their NLCC1 scores are added to generate the weighted score which we term NLCC2. NLCC2 denotes the concentration intensity not only at the point where the local unit locates but it also indicates the manner how that location is neighbored by areas with varied levels in terms of concentration. Put differently, it shows a sort of potentials of respective areas taking externality into consideration. In this sense NLCC2 gives a composite measurement of areal concentration. From the local unit's viewpoint, one can assess with NLCC2 how each local unit locates at a particular subarea which has a wide spectrum of magnitude in terms of concentration.

(3) distribution of scores

(a) NLCC1

Some local units among 14,961 active ones share the same address and thus give identical set of coordinates. By means of spatial aggregation NLCC1 scores are calculated for 10,324 respective points. NLCC1 gives a set of basic statistics: maximum score 77, minimum score 1, mean 1.41, and standard deviation 1.787.

(b) NLCC2

We defined the scope of neighborhood by a specified scale of circle centered by each coordinates. In urban areas coordinates loaded with higher NLCC1 are densely populated, while those with lower NLCC1 may sparsely spread in suburban areas. We can measure how respective points with varied NLCC1 distribute with differed degree of neighborship by assessing the intersection among buffers. Since the NLCC2 score calculated by aggregating the NLCC1 scores spatially also connotes the intensity of concentration in surrounding subareas. We expect, therefore, that NLCC2 can give a good measure to assess local units' concentration in taking externality into consideration.

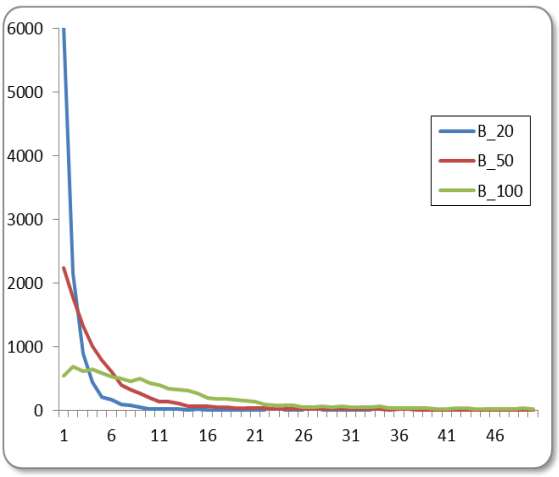
The larger the radiuses are, the more number of locational points are plotted within the buffer and thus the NLCC2 score portrays an amplified picture of local units' spatial distribution. Table 1 shows statistics of the distribution of NLCC2 scores.

Table 1 Distribution statistics of NLCC2

Buffers	Max	Mean	S.D.
20m	77	2.12	2.685
50m	101	6.41	9.606
100m	270	20.05	33.127

As the table indicates respective locational points are loaded with NLCC2 scores ranged from 1 to 77, 101, and 270 for 20m, 50m, and 100m buffers, respectively. Diagram 1 shows histograms of NLCC2 score by size of buffers. It illustrates that larger buffers intersect each other and thus histograms of NLCC2 tell a clear shift while diminishing drastically the score 1 cases.

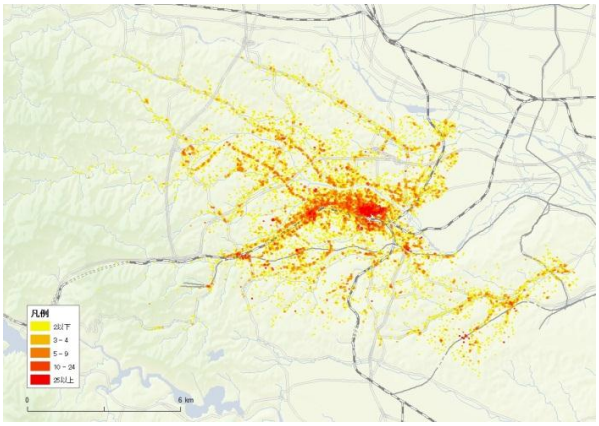
Diagram 1 Histograms of NLCC2 score



6. Conclusions - Spatial distribution of NLCC2 scores

As the NLCC2 gives an indicator of local units' spatial concentration by subareas, one can visualize it on map by delineating zones by degree of NLCC2 by introducing an appropriate size of buffer. Following two gradational maps show how the local units' densely located zones spread within the city by NLCC2 score classes.

Map 1 The whole field of Hachioji



Map 2 The city center area