Spatial-Structural Relations among Technology Industrial Clusters: A Comparative Analysis of Metropolitan Regions in the U.S.

Roger R. Stough*
Rajendra Kulkarni**
Jean Paelinck**

A paper prepared for presentation at the 41st Congress of the European Regional Science Association, Zagreb, Croatia, August 29th-September 1st, 2001
* Corresponding author, The School of Public Policy, George Mason University
[rstough@gmu.edu]
** The School of Public Policy, George Mason University

Abstract

Technical industrial clusters are defined and analyzed for a sample of U.S. metropolitan regions. Economic structure, spatial proximity and shape of the clusters are examined across the metropolitan regions through various economic and spatial measures and statistics. The data for this research are used to test the hypothesis that close spatial proximity results in stronger economic functional interdependence. This hypothesis is examined and the implications of the test are explored.
Introduction.

Industrial clusters are groups of industries that are highly inter-dependent in that they buy and sell from each other, their products tend to be functionally interrelated and there are supporting organizations, e.g., associations, research institutions, etc., related to the cluster functions. As such, the components (e.g., sectors or industries) are usually geographically concentrated in specific regions or in specific parts of states or metropolitan regions. Industrial sectors in the core of a cluster, for the most part, produce for the market outside the local region or area of concentration and therefore tend to be export-base industries. There are different types of industrial clusters including, but not limited to, traditional industry clusters (the dominant industry or group of related industries in a region) to new emergent or propulsive clusters to service-based clusters, etc.

The analysis of industrial clusters has become one of the new major tools used to guide and inform regional development and technology policy. Cluster analyses have recently been conducted at the metropolitan and state level, not just in the U.S., but throughout the world (Rosenfeld, 1997; Glasmeier, and Harrison, 1997; Bergman, Feser and Sweeney, 1996; Bosworth and Brown, 1996; Held, 1996; Jacobs and De Man, 1996; Rosenfeld, 1996; Doeringer, 1995; Saxenian, 1994, 1991; and Porter, 1990). Despite this intense level of activity and historical antecedent work in industrial cluster analysis (Isard, 19xx) this is a relatively new analytical approach and there is no standard methodology. Investigators have utilized a variety of approaches both quantitative and qualitative with the more fruitful studies utilizing both. The quantitative approaches typically analyze industrial sector data using methods that range from measures of industry size and change (e.g., employment, wage level, establishments and related dynamics) to measures of inter-industry linkage levels (e.g., input-output models). Qualitative analysis (interviews, focus groups and surveys) is needed, however, to learn about the structure of supply chains and to evaluate and describe hard and soft infrastructure.

Earlier work by Stough, et al. (2000) developed 15 measures of performance for economic sectors that were used to define and evaluate industrial clusters and their performance. New tools for presenting these structural analyses and results were created including a series of performance indices and spider diagrams for illustrating the results of the analyses. However, none of the many cluster analyses noted above nor the Stough et.
al. study have analyzed the spatial structure of the clusters they have identified. The failure to focus research on this topic identifies a significant problem area because industrial clusters are believed to cluster geographically (although some clusters may be highly diffuse such as the auto industry cluster in some parts of the U.S. like the Southeastern States). More importantly knowing the degree of geographical clustering and the relationship of this to the intensity of inter-sectoral supply relationships could inform development policy.

This paper is an inaugural investigation of the spatial properties of industrial clusters. Several approaches for identifying the spatial clustering of sectors that define industrial clusters are developed and then applied to the analysis of high technology sectors in three metropolitan regions in the U.S. Approaches for measuring the relationship between sectoral clustering and the strength of inter-sectoral dependency, and the geographic shape of industrial clusters are developed. In summary, this paper analyses the spatial structure of a group of hi-technology activities located in three U.S. Metropolitan Statistical Areas (MSA). At the same time as offering insights into the spatial patterns of those activities inside urban conglomerates, new instruments developed to reveal those patterns and to gain insight into their spatial properties. Finally, a test of joint spatial-structural cluster relationships is presented.

The next section presents the database and the overall characteristics of the observed patterns, while sections 3 and 4 dwell on the inter- and intra MSA analyses. Conclusions and references follow as usual.

1. Database.

Table 1 lists the hi-technology sectors selected for this first analysis, to wit 33 SIC four-digit IT (Information and Technology) sectors (Stough et al., 2000; data derived from Business Analyst 1.1, 1999). Sectors were selected based on earlier work by Stough, et. al., 1998.

The MSA’s that have been analysed are:
- the Austin – San Marcos (TX) MSA;
- the Boston (MA and NH) MSA;
- the Washington – MD – VA – WV MSA.
<table>
<thead>
<tr>
<th>SIC Category</th>
<th>SIC</th>
<th>Austin MSA</th>
<th>Boston MSA</th>
<th>Wash MSA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Electronic Computers</td>
<td>3571</td>
<td>34</td>
<td>103</td>
<td>83</td>
</tr>
<tr>
<td>Computer Storage Devices</td>
<td>3572</td>
<td>6</td>
<td>34</td>
<td>15</td>
</tr>
<tr>
<td>Computer Terminals</td>
<td>3575</td>
<td>2</td>
<td>31</td>
<td>18</td>
</tr>
<tr>
<td>Computer Peripherals</td>
<td>3577</td>
<td>33</td>
<td>195</td>
<td>62</td>
</tr>
<tr>
<td>Calculating and Accounting Machines</td>
<td>3578</td>
<td>0</td>
<td>9</td>
<td>7</td>
</tr>
<tr>
<td>Office Machine</td>
<td>3579</td>
<td>2</td>
<td>14</td>
<td>11</td>
</tr>
<tr>
<td>Telephone and Telegraph apparatus</td>
<td>3661</td>
<td>11</td>
<td>78</td>
<td>62</td>
</tr>
<tr>
<td>TV and Cable comm. equipments</td>
<td>3663</td>
<td>14</td>
<td>99</td>
<td>91</td>
</tr>
<tr>
<td>Electron Tubes</td>
<td>3671</td>
<td>1</td>
<td>13</td>
<td>3</td>
</tr>
<tr>
<td>Printed Circuit Boards</td>
<td>3672</td>
<td>32</td>
<td>167</td>
<td>29</td>
</tr>
<tr>
<td>Semiconductor and related devices</td>
<td>3674</td>
<td>49</td>
<td>130</td>
<td>17</td>
</tr>
<tr>
<td>Electronic Capacitors</td>
<td>3675</td>
<td>0</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>Electronic Resistors</td>
<td>3676</td>
<td>1</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>Electronic coils and transformers</td>
<td>3677</td>
<td>1</td>
<td>19</td>
<td>62</td>
</tr>
<tr>
<td>Electronic connectors</td>
<td>3678</td>
<td>3</td>
<td>19</td>
<td>1</td>
</tr>
<tr>
<td>Electronics components, nec</td>
<td>3679</td>
<td>46</td>
<td>219</td>
<td>60</td>
</tr>
<tr>
<td>Magnetic and Optical Recording Media</td>
<td>3695</td>
<td>14</td>
<td>40</td>
<td>36</td>
</tr>
<tr>
<td>Radio Telephone Communications</td>
<td>4812</td>
<td>52</td>
<td>203</td>
<td>217</td>
</tr>
<tr>
<td>Telephone Communications, exc. Radio</td>
<td>4813</td>
<td>206</td>
<td>531</td>
<td>640</td>
</tr>
<tr>
<td>Telegraph and other Communications</td>
<td>4822</td>
<td>12</td>
<td>38</td>
<td>49</td>
</tr>
<tr>
<td>Radio Broadcast Station</td>
<td>4832</td>
<td>44</td>
<td>168</td>
<td>159</td>
</tr>
<tr>
<td>Television Broadcast Station</td>
<td>4833</td>
<td>11</td>
<td>53</td>
<td>91</td>
</tr>
<tr>
<td>Cable and other pay TV services</td>
<td>4841</td>
<td>35</td>
<td>170</td>
<td>118</td>
</tr>
<tr>
<td>Communication services, nec</td>
<td>4899</td>
<td>33</td>
<td>77</td>
<td>138</td>
</tr>
<tr>
<td>Computer Programming Service</td>
<td>7371</td>
<td>560</td>
<td>2157</td>
<td>2049</td>
</tr>
<tr>
<td>Prepackaged Software</td>
<td>7372</td>
<td>216</td>
<td>816</td>
<td>547</td>
</tr>
<tr>
<td>Computer Integrated Systems design</td>
<td>7373</td>
<td>174</td>
<td>806</td>
<td>1236</td>
</tr>
<tr>
<td>Data Processing and Preparation</td>
<td>7374</td>
<td>140</td>
<td>405</td>
<td>561</td>
</tr>
<tr>
<td>Information Retrieval Services</td>
<td>7375</td>
<td>37</td>
<td>120</td>
<td>155</td>
</tr>
<tr>
<td>Computer Facilities Management</td>
<td>7376</td>
<td>4</td>
<td>19</td>
<td>38</td>
</tr>
<tr>
<td>Computer rental and leasing</td>
<td>7377</td>
<td>8</td>
<td>56</td>
<td>45</td>
</tr>
<tr>
<td>Computer maintenance and repair</td>
<td>7378</td>
<td>83</td>
<td>342</td>
<td>379</td>
</tr>
<tr>
<td>Computer related services, nec</td>
<td>7379</td>
<td>272</td>
<td>1174</td>
<td>1523</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td>2141</td>
<td>8313</td>
<td>8503</td>
</tr>
</tbody>
</table>

\[
t_k = n_k/n
\]  

(1)

They are designated by A, B and W respectively. Table 2 provides the data for the number of plants observed in each MSA; they total 2,141 (A), 8,313 (B) and 8,503 (W) respectively. The data for this study include plants by information technology sector (SIC number) and the respective geographical coordinates (source: own computations); Maps 1, 2 and 3 reproduce that information graphically.

[ Map 1, 2 and 3 about here…]

The subsequent parts of the paper present a more analytical description of the observed patterns.

3. Inter-MSA comparisons.

The first coefficient to be computed is what can be termed a Tinbergen-coefficient; it is derived from Tinbergen-Bos spatial economic equilibrium analysis in terms of “centers” and “systems” (Paelinck, 2000), centers being defined as spatial clusters of activities, systems as spatial combinations of centers.
The Tinbergen coefficient is defined as the relative number of sectors present in an observed center, i.e.:

\[
\frac{n}{n_k}
\]

where, \( n \) is the total number of sectors analysed (in casu 33) and \( n_k \) the number of sectors effectively observed in a given center \( k \) (\( k=1,2,3 \)).

Table 1. Information and Technology sectors by Regions

For A, B, and W respectively the \( t_k \)’s were .9394, 1 and .9697, with only 2 sectors being absent in A and 1 in W (see Table 1).

The plants’ density by population and by area were also computed for each technology sector. The results are as follows: the number of plants per 100 thousand population are 187 (A); 140 (B) and 179 (W); the number of plants per square mile are 0.50 (A); 1.23 (B) and 1.29 (W); showing the effects of different center sizes and population densities.

The average number of plants per sector is 65 (A), 252 (B) and 258 (W), with coefficients of variation (standard deviations divided by the respective means) of 1.7481, 1.7337 and 1.8408.

In these global terms, and taking into account the standardizing deflators (population and surface) the results point at a certain, though not complete, degree of homogeneity in the general (still not spatial) patterns observed.

This relative homogeneity is confirmed by the matrix of correlation coefficients and its eigenvalues; still in the ABW-order the correlation matrix is \( \{1, .9883, .9485; 1, 9648; 1\} \) with eigenvalues of 2.9345, .0558 and .0097; it is known that if \( n-1 \) eigenvalues out of \( n \) are near zero, the overall correlations are extremely high (positively or negatively, but in the present case positively as the simple correlation coefficients show). A measure of the overall correspondence might be the largest eigenvalue divided by the sum of the eigenvalues, in casu \( 2.9345/3=.9786 \). Figures 1, 2 and 3 reproduce those observations graphically.

[Figure 1, 2 and 3 about here…]

4. Intra-MSA analyses.
Spatial analysis requires the introduction of topological elements; these are now introduced in terms of relative positions (coordinates) and distances; the distances have been defined as Manhattan distances (sum of the absolute differences of the respective x and y coordinates), a rather realistic metric for the study of urbanized areas.


A first indicator of the intra-MSA spatial structure is the average distance (total distance depending on the number of plants) separating the plants analyzed, divided by the square root of the metropolitan area in square miles (this to ensure dimensional homogeneity of the numerator and denominator); the resulting A, B an W indicators are respectively 0.0437, 0.0692 and 0.0727, showing different orders of magnitude of mutual internal relative accessibility. Average distances are 0.0446, 0.0952 and 0.0977, confirming the previous observation.

Returning to graphs 1, 2 and 3, one can visualize the (unweighted) centers of gravity of the hi-tech activities present; noticeable are the differences in shape of the spread of those centers, a fact which is submitted to further mathematical analysis.

To better understand these spatial linkings, the following approaches have been envisaged:

- compute the Hausdorff distances (Hausdorff, 1962, pp.166 ff.) between all the plants belonging to different activities; this allows to compare the relative closeness of the sectors involved, and to examine the hypothesis that more centrally clustered sectors have higher input (supply chain) dependencies (measured, e.g., by the sum of the corresponding input coefficients) through correlation analysis;
- conduct a nearest neighbor analysis in terms of the average nearest neighbor distances between plants belonging to different sectors, and apply again the above analysis of the relationship to the input coefficients.

As an intermediate investigation, the distances between the sectoral centers of gravity referred to above were used, together with aggregated summed input coefficients. (taken from Survey of Current Business, 2000); table 3 hereafter shows the aggregation.
Table 2: aggregation of sectors

<table>
<thead>
<tr>
<th>Input-output code</th>
<th>Sector</th>
<th>SIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>51</td>
<td>Computer and office equipment</td>
<td>357</td>
</tr>
<tr>
<td>56</td>
<td>Audio, video and communication equipment</td>
<td>365-366</td>
</tr>
<tr>
<td>57</td>
<td>Electrical components and accessories</td>
<td>367</td>
</tr>
<tr>
<td>66</td>
<td>Communications, except radio and TV</td>
<td>481, 482, 484, 489</td>
</tr>
<tr>
<td>73A</td>
<td>Computer and data processing services</td>
<td>737</td>
</tr>
</tbody>
</table>

The simple correlation coefficients between distances and the summed input coefficients were \(-0.5008\) (A), \(-0.7335\) (B) and \(-0.3890\) (W), showing all of them to be negative relations between distances and summed input coefficients, as expected. The strongest relation was observed in the Boston area. Once more, this is only an intermediate investigation; additional analyses will be required at much more disaggregated levels.

4.2. Dispersion, Orientation and the shape of the distribution of technology companies

The presentation of spatial analysis and the results computed below are based on CrimeStat (1.0) from the National Institute of Justice (U.S. Dept. of Justice, U.S. Government, 2000.)

Dispersion around the mean center of a region

Standard distance measures the average of distances between companies and the mean center of a region. The mean center of a region is its geographic centroid. It is computed as follows:

\[
d_{xy} = \text{standard distance of a distribution} = \sqrt{\frac{\sum_{i} (d_{i,\text{mean center}})^2}{n-2}}
\]  

where, \(d_{i,\text{mean center}}\) is the distance between company \(i\) and the mean center of a region and \(n\) is the total number of companies in a region. The Standard distances for A, B and W are respectively, 10.67 miles, 23.19 miles and 16.46 miles and this suggests that region A is more tightly clustered than region W, which in turn is more clustered than region B.

Shape and orientation
So far, we have presented dispersion of companies across a region and their concentration around the mean center. Next, we look at the shape and orientation of the spread of technology firms in each region. The standard deviational ellipse is a measure skewness of the distribution of technology companies. It is computed as follows:

\[
Std. \text{ Deviational Ellipse } (DSE) = \sqrt{\frac{\sigma_x^2 + \sigma_y^2}{2}} 
\]  

(3)

where \(\sigma_x\) and \(\sigma_y\) are standard deviations along \(X\) and \(Y\) directions with \(X\) and \(Y\) being orthogonal to one another. These two are perpendicular to each other and hence they define an ellipse. Figures 4, 5 and 6 show one and two SD standard deviational ellipses for each of the regions. The following table summarizes the statistics for the major and minor axes of the ellipses.

The ratio of the ellipse axes for the three technology establishment distributions are: Austin (0.61), Boston (1.48) and Wash. DC (1.06) suggest that, the Washington DC region has the most symmetrical (nearly circular) distribution of technology companies, while both Austin and Boston have skewed distribution that is opposite of each other.

<table>
<thead>
<tr>
<th>Table 3. Elliptical distribution</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Length in miles (1 SD)</td>
</tr>
<tr>
<td>Length in miles along axis for 1 SD</td>
</tr>
</tbody>
</table>

The ratio of the ellipse axes for the three technology establishment distributions are: Austin (0.61), Boston (1.48) and Wash. DC (1.06) suggest that, the Washington DC region has the most symmetrical (nearly circular) distribution of technology companies, while both Austin and Boston have skewed distribution that is opposite of each other.

The above computations for ellipses are carried by rotating \(X\) and \(Y\) axes such that the sum of squared of distances between company locations and the axes are minimized.
Angle of rotation and hence orientation for Boston (84.63) is the highest, while that for Washington is 71 degrees and Austin is just 12.9 degrees.

[Map 4, 5 and 6 about here…]

These results show the differences for the three areas. It is interesting to note that, Wash. DC region shows the most symmetrical distribution, while Boston shows the largest stretch (1.48) and tilt (angle of rotation) (84.6 degrees), and the Austin area shows the least stretch (0.61) and tilt (just 12 degrees).

4.3. Nearest Neighbor and L-statistic analyses of the Washington D.C., PMSA

Next, we used the Nearest-neighbor index and L-statistic (also known as Ripley’s K statistic) to examine the distribution of technology industry sectors (See table 4) in the Washington DC PMSA. The Nearest-neighbor index (NnbrI) helps describe the pattern of spatially distributed points over a study area, eg., whether such a pattern is clustered, random or dispersed. In the case of random patterns, each location is equally likely to be occupied by a point and thus such point locations are independent of each other.

Table 4. Selected Technology sectors in the Washington D.C., PMSA

<table>
<thead>
<tr>
<th>SIC</th>
<th>Technology sector</th>
<th>Number of companies</th>
</tr>
</thead>
<tbody>
<tr>
<td>7371</td>
<td>Computer Programming services</td>
<td>2,049</td>
</tr>
<tr>
<td>7372</td>
<td>Prepackaged software</td>
<td>547</td>
</tr>
<tr>
<td>7373</td>
<td>Computer integrated systems design</td>
<td>1,048</td>
</tr>
<tr>
<td>7374</td>
<td>Data processing and preparation</td>
<td>561</td>
</tr>
<tr>
<td>7375</td>
<td>Information retrieval services</td>
<td>155</td>
</tr>
<tr>
<td>7376</td>
<td>Computer facilities management</td>
<td>35</td>
</tr>
<tr>
<td>7377</td>
<td>Computer rental and leasing</td>
<td>48</td>
</tr>
<tr>
<td>7378</td>
<td>Computer maintenance and repair</td>
<td>369</td>
</tr>
<tr>
<td>7379</td>
<td>Computer-related services</td>
<td>1,523</td>
</tr>
<tr>
<td>8711</td>
<td>Engineering services</td>
<td>2,113</td>
</tr>
<tr>
<td>8742</td>
<td>Management services</td>
<td>1,041</td>
</tr>
<tr>
<td>8743</td>
<td>Management consulting services</td>
<td>6,224</td>
</tr>
</tbody>
</table>
On the other hand, non-random patterns occur when the locations of points are
dependent on each other. There are two types of non-random patterns, 1) clustered
patterns indicating an attraction for the phenomenon to locate proximally and 2) dispersed
patterns that indicate a repelling property that makes points locate as far away from each
other as possible.

Consider a study area A with a set of spatially distributed points (N). Then NNbr
index is computed as the ratio of the average or mean distance between N points and the
expected mean distance if these points are randomly distributed.

Average distance between nearest neighbors
\[d_{NNbr} = \frac{\sum_{i=1}^{N} \min(d_{ij})}{N}\] (4)

where, \(\min(d_{ij})\) is the minimum of distances between a given point \(i\) and all other points
\(j \neq i\). Let \(d_{rand}\) be the expected mean random distance for N points distributed over area
A. The \(d_{rand}\) is defined as:

\[d_{rand} = \frac{1}{2\sqrt{\frac{A}{N}}}\] (5)

Then, NNbrI is given by:

\[NNbrI = \frac{d_{NNbr}}{d_{rand}}\] (6)

Thus, NNbrI is 1 when the observed mean distance and expected mean random
distance is same. On the other hand, a NNbr Index value of less than 1, indicates clustering
and more than 1 indicates dispersion. Note that the NNbrI can be computed for first
nearest, 2\(^{nd}\) nearest, 3\(^{rd}\) nearest, \ldots n\(^{th}\) nearest neighbors and so on. The following chart
(Figure 4) shows the results for 25 nearest neighbors for most of the technology sectors
(except for two sectors 7176, Computer facilities management and 7177, computer rental
and leasing, dropped because of small n). It appears that all the technology sectors show
clustered patterns and nearly all the technology sectors approach a “steady state” after the
10\(^{th}\) nearest neighbor. Among the technology sectors for which this analysis was carried
out, management consulting services (8742) shows the most clustered pattern and the
information retrieval services (7375), the least.

[Figure 4 about here…]
Next we computed L statistic for the same set of technology sectors. The L Statistic is a non-randomness statistic for spatially distributed point data. It is also known, as Ripley’s K statistic. It provides a spatial test for non-randomness for distances that range from very small to large covering the entire study area. For example, consider a study area A with N points. For each of the points one can draw a circle of radius r and count the number of points inside that circle. In case of a random distribution of N points within an area A, the number of points per unit area is: 

$$P = \frac{N}{A} \quad (7)$$

Then for an area with radius r the expected number of points is:

$$P_e = P \left( \pi r^2 \right) \quad (8)$$

For a specific case, the actual number of points could be more (indicating clustering) or less (dispersion). The L-statistic is computed as follows: For each of the points in a given study region, count the number of points within a radius r using equation (8). Repeat this for every point in the study region. Next compute the average over all the points to complete the K statistic as follows:

$$K(r) = \frac{1}{PN} \sum_{i} \sum_{j \neq i} \left( \text{# of points in an area with radius r at point } i \right) \quad (9)$$

And the L statistic is given by:

$$L(r) = \sqrt{\frac{K(r)}{\pi}} - r \quad (10).$$

One may repeat this computation for increasing values of r. Using equation (7) through (10), the L-statistic was computed for all the technology sectors in the Washington DC PMSA. A plot of L statistic against distance is shown in Figure (5) for the technology sectors. All plots have inverted u shape, indicating that each sector shows clustering at some distance from the geographic mean and dispersion afterwards. Among the technology sectors, again the management consulting services sector (8742) shows a sharp increase in clustering up to 20 miles and dispersion at longer distances. While, the information retrieval services (7375) shows a broader clustering and dispersion pattern.

[ Figure 5 about here…]

Next, we constructed a vector of distances (d_{ij}) between mean latitude and longitude coordinates between each pair of technology sectors. Similar computations were
carried out for both the NnbrI and L-statistic. The following table shows the values of correlation coefficients between the three vectors, latitude-longitude distances, the NnbrI and L-statistic. It was expected that there would be a strong correlation among all these distance statistics. Instead, (surprisingly) there is a very small positive correlation between the L-statistic and $d_{\|}$, while the other two show negative correlation.

Table 6: Correlation Coefficient the Washington D.C., PMSA

<table>
<thead>
<tr>
<th>Correlation Coefficient Between L-statistic and $d_{|}$</th>
<th>Correlation Coefficient between L-statistic and NnbrI</th>
<th>Correlation Coefficient between NnbrI and $d_{|}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.23377</td>
<td>-0.182978</td>
<td>-0.24016</td>
</tr>
</tbody>
</table>

5. Extensions.

Apart from the Hausdorff distance and the nearest neighbor analyses mentioned earlier, a cluster analysis of the individual plant locations is envisaged. The following specification has been chosen:

$$\min \varphi = \sum_{q<p} x_{pq} d_{pq} x_{qc}$$

$$x_{pc}$$

s.t.:

$$x_{11} = 1$$

$$\sum c x_{pc} = c^*, \forall p$$

$p$ and $q$ are plant indices, $c$ a cluster index, and $d_{pq}$ distance condition (6) is a binary condition, but if it is relaxed to $0 < x < 1$, fuzzy clustering may result. One can see that in fact one maximizes internal cohesion or interaction; if in (4) production levels or employment are to be integrated, their inverse products should be used. One can also restrict interaction to activities of a different nature.

An example to illustrate this: take four plants located at distances \{5,10,15;7,13,5\} to be clustered into two clusters ($c^* = 2$); function (4) then becomes:

$$\varphi = 1*(5x_{21} + 10x_{31} + 15x_{41}) + 7*(x_{21}x_{31} + x_{22}x_{32}) + 13*(x_{21}x_{41} + x_{22}x_{42}) + 5*(x_{31}x_{41} + x_{32}x_{42})$$
The solution is \( x_{21} = x_{12} = x_{42} = 1 \), giving \( q = 10 \), this function is obviously non-increasing for increasing \( c^* \)

6. Conclusions.

This paper has made an initial investigation of the spatial properties of the location of plants in multiple industrial high technology sectors. It has demonstrated various measures of spatial clustering and analytical techniques that explore the relationship between spatial closeness and functional interrelationships, such as potential supply chain relationships. Next, it developed and demonstrated a measure of the geographic shape of cluster distributions. It is important to note that the variations in the geographic shape of the distributions seemed in all three cases to be dependent on the physical road infrastructure. After determining the shape and orientation of the distribution of technology companies, a multi order nearest neighbor index (NnbrI) was computed for the nearest 25 neighbors for a selected number of technology sectors in one of the technology regions (W or Washington D.C. PMSA). The multi order NnbrI index settles down to a steady value after about 10\(^{th}\) nearest neighbor. It suggests that almost all of these sectors are well defined clustered within a small area of the Washington D.C., PMSA. Next, the L-statistic was used to further analyze the clustering/dispersion for each of the selected technology sectors in the Washington D.C., PMSA region. Again, the L-statistic analysis confirms that nearly all sectors have clustering tendency within a 25 mile radius from the geographic mean. Once this distance threshold is crossed, all these sectors show a tendency towards dispersion of the technology companies. Also, computed was a correlation coefficient between the following measures

1. Distances between mean or average of each sector
2. The Nearest neighbor statistic up to the 10\(^{th}\) nearest neighbor
3. The distances associated with the L-statistic maxima.

These correlations are weak and do not shed any new light on the distribution of the technology companies in the region. A more careful analyses is planned for the future and will be carried out for all the three regions.

These measures all hold the potential for advising regional economic development and technology investment policy.
7. References


Business Analyst 1.1, 1999, Extension to ArcView 3.2a (ESRI).


***

Map 1. Austin-San Marcos (TX) MSA
Map 2. Boston (MA, NH) MSA

Map 3. Wash D.C. (Wash, D.C., VA-MD-WV) MSA
Figure 3. Wash, D.C. vis. Austin MSA: Information Technology and Telecom Sector Plants

Map 4. Austin MSA
Map 5. Boston MSAMap

6. Washington D.C., PMSA
Figure 4. Near-neighbor Index (NnbrI) by sector in the Washington D.C., PMSA

Figure 5. L-statistic by sector in the Washington D.C., PMSA