Disparities in Distribution of Information Resources in the City of Philadelphia: Spatial Analysis of Socio-demographic and Information Resource Variables

Emad Khazraee  
Drexel University  
emad@drexel.edu

Adam Townes  
Drexel University  
amt74@drexel.edu

Anthony Grubesic  
Drexel University  
ahg32@drexel.edu

Abstract

This study explored the use of GIS techniques to spatially explore and identify areas where provision of public library and information services are insufficient. By identifying the spatial distribution of inequity of access to information services in Philadelphia, Geographic analysis techniques were applied to investigate the existence of a relationship between spatial orientation, socio-economic factors and information access. The findings of this study suggest the existence of a spatial pattern in relation to disparity of the distribution of information resources in Philadelphia. A secondary analysis utilized the aforementioned findings to assess the existence of patterns of library service provision on a limited data set. The results of analyses indicate that in the disadvantaged areas of Philadelphia, public libraries play the role of information access gateways/bridges. The findings can be used to better locate the future public library and information service centers in the city of Philadelphia.

Keywords: information equity, information access, GIS, geographic information system, urban information gateways

Introduction

The goal of this research is to study the spatial patterns of access to information resources in Philadelphia and the relation of those patterns to the spatial distribution of social and economical factors to suggest the best location for the new information service centers in the city of Philadelphia in response to identified areas of need and inequality.

Information inequality has been an issue of great concern and study for at least 50 years (Yu, 2006). In an historical sense, research into information inequality seemingly coincides with the rise of a postindustrial information society, the civil rights movement and the federal government’s War on Poverty during the 1960’s. Research from late 1960’s through the 1980’s focused on understanding the interaction between information need and some form of disadvantage. Disadvantage was characterized by lack of knowledge in regard to formal information channels, frequent television viewing, sparse usage of reading materials and association with informal information networks that were deficient in the information resources and channels (Childers and Post, 1975; Yu, 2006).

It was during the 1990’s that the concept of a digital divide came to prominence because of the apparent social, economic and racial differences in ownership of home computers and home based Internet connections. More recently topics such as the app gap and the time-wasting gap (Richtel, 2012) indicate that even where disadvantaged populations do have access, pre-existing social, educational and economic problems still influence information seeking behavior and need.

Information behavior is a complex and multifaceted concept. Issues such as inequitable distribution of information resources, lower educational achievement and lack of digital literacy can affect

Acknowledgements: We would like to thank the staff of the Free Library of Philadelphia for their help in providing library statistics and qualitative information in regard to patron usage patterns. Particular thanks to Jeff Bullard, Tamikka Coppin, Chris Kottcamp, Edwin Saeger and Sandra Thompson.


Copyright is held by the authors.
the information behavior of underserved populations. In order to better understand the information needs and policy considerations related to provision of information resources, information access and library services, it is critical to recognize the demographic, social and geographical elements which relate information inequity. This study utilizes a geographic information system to explore the socio-economic and demographic factors of underserved populations in Philadelphia and how they relate to the imbalance in spatial orientation of information resources in the city.

**Literature Review**

The essential point to be emphasized is that information inequality entails more than simply technologically, socially, economically or institutionally determinant issues, but a combination thereof. Grubesic and Murray (2002) recognized that the digital divide is more dynamic and intricate than many studies suggest. Moreover, it is not simply confined to areas on the basis of socio-economic determinant factors, but is also spatially defined. The spatial orientation of such factors can be tied to a user or a group is useful when describing information behaviors of disadvantaged populations. Information use environments can be understood to be the set of elements which affect the availability, access to and use of information (Taylor, 1991).

As it relates to information behavior, lack of information access is an institutional impediment to which disadvantaged populations adapt. Essentially there is an interaction between infrastructural factors such as level of broadband provision or distance to nearest library and information seeking behavior. Sin and Kim (2008) as well as D’Elia (1980) found significant associations between distance from an information resource, such as the public library and use or nonuse of that resource.

The intent of this study was to explore possible correlative relationships between access to information resources and socio-demographic variables that interact at the neighborhood and community level in Philadelphia. GIS analysis provides a powerful tool to graphically and spatially orient the coalescence of infrastructural, environmental and demographic factors. The spatial analysis performed in this study helps to define information use environments in the various neighborhoods and areas of Philadelphia. Spatial analysis is useful for understanding the geographical context of information access issues. For instance one simple and easily mapped factor that can be tied to use is distance to nearest library. Sin and Kim (2008) utilized GIS techniques and found that distance to the nearest public library branch is a statistically significant factor in determining library use by communities. The location of the nearest branch matters in cases where users are not able to afford Internet, or live in areas with low provision.

Koontz et al. (2009) utilized GIS to analyze demographic and socio-economic variables of service areas of closed library facilities in a nationwide survey. The results from Koontz et al. (2009) indicate that the majority of public library closures occur in areas with lower levels of income and education. Consider then, that the presence of a library in a community may represent an enriching factor in a neighborhood information environment. Accordingly, Sin (2010) concluded that the neighborhood information environment can affect information use across individual, household and community levels. By visualizing the existing patterns of information inequity in relation to heterogeneous geographical distributions of wealth, imbalance in distribution of information resources, and socio-demographic factors (e.g. unemployment, household income, and education level) a deeper understanding of communal behavioral norms can be gained. Such an understanding is useful in enacting effective policies and solutions to address information inequity. Understanding the localized problems and association information needs of populations in the underserved areas allows for design and implementation of information resources and services that are relevant and sensitive to communal social norms and needs. Further, development of research on information behavior in relation to the convergence of factors that affect information behavior, within a geographical context has the potential to enhance understanding of general information of underserved populations.
Data Description

Demographic, economic, educational, crime and information access data were collected according to geographic delineations, such as census block groups, census tracts and breakdown by neighborhood. Broadband provision density and sum total of broadband accounts in each block group were also included in the analysis. The data was acquired from various data sources such as the US Census Bureau, Philadelphia NIS CrimeBase, PASDA, Free Library of Philadelphia and NTIA. For this study demographic data (population density and proportion of White or African American population), economic factors (median family income, values of houses, median monthly rent, percentage of vacant properties, and percentage of vacant residential buildings), educational factors (total number of higher education attainment and based on gender) were acquired from the US Census\(^1\) 2000 data. Data for the crime (aggregated assaults, all serious assaults numbers and assault rate per 1000 population) was collected from the Philadelphia NIS CrimeBase\(^2\). The information about geographic distribution of public libraries was provided by Pennsylvania Spatial Data Access (PASDA)\(^3\). The Free Library of Philadelphia provided the monthly statistics of turnstile, circulation and computer use for nine branches for one year. Finally, broadband provision density and sum total of broadband accounts in each block group of city of Philadelphia in 2009 were acquired from National Telecommunication and Information Association\(^4\) data.

Methods

The goal of this research was to understand the spatial pattern in the distribution of information resources and that distribution’s relation to the spatial orientation of other social and economic factors. Two sets of analyses were performed and are described in the following paragraphs.

Tobler’s First Law of Geography indicates that ”everything is related to everything else, but near things are more related than distant things“ (Tobler, 1979, 236). In terms of analysis we first conducted an exploratory analysis of the different distributions of demographic, economic, educational, crime and information access data in the Philadelphia area. In this step distribution maps were created to compare the contrasts in different neighborhoods. Then correlation tests were carried out to assess if any of the included socio-economic variables were correlated with each other. Based on exploration and analysis of the geographic distribution of different factors in the city of Philadelphia, the most significant variables were identified for incorporation into the spatial cluster analysis. These variables are Median Family Income, Population Density, Crime Rate, Higher Education and Race (percentage of White or African American population). Spatial cluster analysis is the spatial grouping of geographical units based on their proximity and their characteristics (Jacquez, 2008). “Spatial distributions with values at certain locations showing relationship with values at other locations are named spatial autocorrelation. Spatial cluster is positive spatial autocorrelation when similar values are spatially clustered together” (Lu, 2000). Spatial cluster analysis was utilized to examine whether urban clusters can be defined based on socio-economic factors and the access rate to broadband Internet connections. Spatial cluster analysis was performed with consideration of the spatial proximity of census block groups in relation to major socio-demographic factors mentioned above.

The secondary analysis was more preliminary in nature, and explored the usage rates of nine public libraries in order to assess whether different patterns of use did exist across branches and what the pattern meant when considered in the spatial context of the spatial cluster analysis. Locations of public libraries were also examined in regard to general coverage, to recognize the undeserved areas. Considerations of coverage patterns were based upon the convenient walking distance and speed, which was assessed as a half mile\(^5\).

---

\(^2\) [http://cml.upenn.edu/crimebase/](http://cml.upenn.edu/crimebase/)
\(^3\) [http://www.pasda.psu.edu/](http://www.pasda.psu.edu/)
\(^5\) Leon Krier who is a pioneer in urban design defines a convenient walking distance as a 10 minute walk (Krier, 2009). We interpret this notion as a half mile distance. The same metric was used recently by the Philadelphia Research Initiative of the PEW charitable trust to evaluate the coverage of the Free library branches in Philadelphia.
Results and Discussion

The results of this study indicated that Philadelphia is spatially heterogeneous in regard to the distribution of race, income and education level, as well as some disparity in broadband access. The spatial cluster analysis showed that the city of Philadelphia is clustered into five areas. Table 1 represents the socio-economic characteristics and average internet access rates for census blocks in each cluster.

Table 1
Summary of spatial cluster analysis

<table>
<thead>
<tr>
<th>Cluster</th>
<th>Median Family Income</th>
<th>Total Higher Education Degrees</th>
<th>Aggregated Assault Rate</th>
<th>Sum Count of Internet Subscriptions</th>
<th>Percentage of African American Population</th>
<th>Percentage of White Population</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>117972</td>
<td>444</td>
<td>14.83</td>
<td>49</td>
<td>15%</td>
<td>76%</td>
</tr>
<tr>
<td>2</td>
<td>25308</td>
<td>85</td>
<td>8.52</td>
<td>35</td>
<td>57%</td>
<td>25%</td>
</tr>
<tr>
<td>3</td>
<td>48119</td>
<td>248</td>
<td>3.71</td>
<td>46</td>
<td>9%</td>
<td>82%</td>
</tr>
<tr>
<td>4</td>
<td>30919</td>
<td>138</td>
<td>8.51</td>
<td>34</td>
<td>67%</td>
<td>25%</td>
</tr>
<tr>
<td>5</td>
<td>60363</td>
<td>327</td>
<td>3.75</td>
<td>38</td>
<td>45%</td>
<td>49%</td>
</tr>
</tbody>
</table>

Figure 1 graphically presents the spatial orientation of the five clusters. Cluster 1 corresponds to areas of very high income and high levels of access. Cluster 2 represents a very densely populated, low income low access area. Cluster 3 corresponds to less populated areas of Philadelphia with high income and high access, but not as high as Cluster 1 or Cluster 5. Cluster 4 represents high access, low income, fairly dense population. Cluster 4 probably represents an area with a large population of college students and middle income households. Cluster 5 is composed relatively high income households with high access. Cluster 5 may be constituted of upper middle class households.

When considering Median Family Income, Population Density, Crime Rate, Higher Education and Race (White or African American), Philadelphia has a population distribution that is racially related, at least where spatial orientation, economic status and demographics variables are considered. The results from cluster analysis demonstrated that levels of education and wealth tend to concentrate geographically in particular areas of the Philadelphia. Though not necessarily correlated, these factors did tend to cluster along racial lines and appeared to be based on levels of income and education. Income and education were somewhat positively correlated. The identified concentrations also applied to the intensity of broadband provision and sum count of broadband subscriptions. The areas of actual usage were more closely positioned to the wealthier and more highly educated regions of the city.

The results of the secondary analysis provide findings related to library services. Based on radial half-mile coverage zones of library service it was determined that there is a gap in library coverage in a densely populated and underserved area as indicated in Figure 2 by the shaded circle with the dotted outline. The area around the library (Nicetown Branch) is underserved and highly populated by low income African Americans, with generally low levels of educational achievement. These findings confirm Koontz et al.’s (2009) identification of fewer libraries in poor areas and Sin and Kim’s (2008) findings related to access and use of library services by disadvantaged populations.

Our library usage analysis at this stage was limited to nine branches because at the time of this study just nine branches data was available (library use statistics refer to year 2010; Table 2). Therefore, this analysis is more exploratory in nature to learn whether it can lead to usable findings. Initial results indicate that two of the nine branches (Lehigh Avenue and Haddington branches) act as information gateways. This finding is based on the comparison of turnstile count versus total computer use. In these two branches, the computer use in relation to turnstile is high in comparison with other branches. Information gateway branches are situated on the border between relatively high information access areas and relatively low information access areas, identified in the spatial cluster analysis. Burghardt’s (1971) hypothesis of gateway cities describes a similar phenomenon where people residing in areas adjacent to the divisions between two geographic units, such as neighborhoods and can be associated with access to some resource such as an oasis in an arid area or areas of differing production such as small cities situated on the periphery of rural areas and large metropolitan areas.
The gateway hypothesis is applied here in order to interpret findings regarding libraries situated adjacent to areas of low and high information provision. Residents of the relatively low provision area, utilize the branches in the high access area as a means of bypassing the disparity of access in their own area. Interpretation of the findings based upon library buffer zone and usage analyses verifies the spatial pattern recognized for the digital divide in Philadelphia from the first analysis.

We follow a cross-sectional analysis method for this exploratory study. This study utilizes data set from the Census year 2000 and broadband access in 2009. For future work we will examine spatial patterns of information resource distributions with longitudinal datasets (US Census 2000 and 2010, as well as longitudinal Internet access data) to better understand the dynamics and patterns of information resource distribution in Philadelphia.

Figure 1. Spatial Cluster Analysis of Philadelphia
Figure 2. Half Mile Radius Library Buffer Zone

Table 2
Summary of the library use statistics of the nine branches of Free Library of Philadelphia

<table>
<thead>
<tr>
<th>Branch</th>
<th>Turnstile</th>
<th>Circulation</th>
<th>Computer Use</th>
<th>Computer use percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eastwick</td>
<td>34552</td>
<td>53822</td>
<td>9101</td>
<td>0.26</td>
</tr>
<tr>
<td>Haddington</td>
<td>38996</td>
<td>31507</td>
<td>13689</td>
<td>0.35</td>
</tr>
<tr>
<td>Haverford Avenue</td>
<td>68500</td>
<td>37834</td>
<td>19856</td>
<td>0.29</td>
</tr>
<tr>
<td>Kensington</td>
<td>52669</td>
<td>34514</td>
<td>16523</td>
<td>0.31</td>
</tr>
<tr>
<td>Kingsessing</td>
<td>89374</td>
<td>42948</td>
<td>16069</td>
<td>0.18</td>
</tr>
<tr>
<td>Lehigh Avenue</td>
<td>43997</td>
<td>28434</td>
<td>22889</td>
<td>0.52</td>
</tr>
<tr>
<td>Regional - West Philadelphia</td>
<td>226616</td>
<td>106034</td>
<td>40307</td>
<td>0.18</td>
</tr>
<tr>
<td>Walnut Street</td>
<td>106453</td>
<td>168929</td>
<td>20082</td>
<td>0.19</td>
</tr>
<tr>
<td>Widener</td>
<td>84351</td>
<td>37078</td>
<td>16261</td>
<td>0.19</td>
</tr>
</tbody>
</table>
**Limitations**

There are two limitations to this study. The first is the limited data from nine of the fifty-four libraries in the Free Library of Philadelphia systems. To verify our findings we need to utilize the complete statistics of all branches. The second limitation, our data about Internet access were limited to broadband accounts and it did not include data about wireless and mobile access. To better describe information poverty we need to use both data about broadband and wireless/mobile access. Recent studies suggest that specific populations tend to access internet through cell phones (Horrigan, 2009). We aim to address these limitations in our future work. Also, it is important to include additional information resources such as community centers, healthcare centers, schools and churches. However, this requires an evaluation of the importance of these centers to access to information.

**Conclusion**

In brief, the major findings of this study were that the distribution of information resources in Philadelphia are spatially heterogeneous and are correlated with spatial distributions of income level, population density, crime rate, education and race, as well as some disparity in broadband access and library coverage. Spatial cluster analysis shows that city of Philadelphia clustered into five spatial clusters regarding these variables. This study shows that GIS techniques can be used to identify underserved populations in urban areas according to the spatial distribution of socio-demographic and geographic variables. Additionally, this study suggests that two public library branches of The Free Library of Philadelphia, (Haddington and Lehigh Avenue branches) appear to act as information resource gateways. Effectively these branches are located near the spatially defined boundary of high and low information access areas. These branches act as information access entry points for people who live in underserved areas. Branches near these spatially defined boundaries tended to exhibit comparatively higher computer usage statistics in comparison to other branches located in others parts of the city.

Utilization of GIS techniques to understand the provision of public library services can assist in ensuring effective utilization of limited resources. Identifying the correlation of the geographical concentration of socio-demographic factors and the location of information resources has the potential to engender a broader understanding of why historically or traditionally underserved populations have in many cases remained underserved in the contemporary era.

**References**


