GEOGRAPHIC INFORMATION DESIGN

for Envisioning and Choosing Neighborhoods

Earl G. Bossard

Abstract: An effective design layout of geo-demographic information is a key component to understanding and ultimately selecting neighborhoods that meet specific requirements. GIS is leading the way to our understanding places, or envisioning them, by integrating dynamic maps and graphics that together display far more than just census or other static geo-demographic data by adding context. Choosing a neighborhood in which to work, shop, reside, or recreate is now possible in the virtual reality of GIS hyperspace.

This paper demonstrates how GIS can be at the heart of a framework of powerful envisioning techniques to ensure enlightened choices. Organized collections of small multiple maps, charts, dynamic images, and data tables, called schemas, are used to facilitate choosing neighborhoods within a region that fit a particular set of perceived needs. Our test search for "Urban Hip Professional" neighborhoods uses census and other geo-demographic data, such as ESRI's Community Tapestry data, placed in logical context across space, scale, and conditions to find neighborhoods that meet specific requirements.

Keywords: data filtering, data queries, data synthesis, envisioning, geographic information systems, information design

GEOGRAPHIC INFORMATION DESIGN

FOR ENVISIONING AND CHOOSING NEIGHBORHOODS

1 INTRODUCTION

This paper demonstrates alternative procedures for using digital data to choose or find spatial locations which must meet multiple criteria. The digital information revolution over the past several decades has created vast sources of information and powerful capabilities for processing that information so that information searchers with modest computer, internet access, and statistical skills have many choices for pursuing searches.

Typical solutions entail using database queries and GIS overlay maps to find locations satisfying multiple criteria. What we shall call the "Simple Yes-No Query/Overlay Technique" often works well to expediently find acceptable locations. Sometimes however, this technique fails to obtain satisfactory results for the reasons outlined in table 1.

Table 1: Reasons for Simple Yes-No Query/Overlay Technique Failure

- Too many solutions may be found, overwhelming the analyst
- Too few, or no solutions may be found, leading the analyst to consider alternatives
- The spatial pattern of the solutions may be unsatisfactory:

  Too clustered
  Too dispersed
  Not related well to areas with special needs

One solution can be using fuzzy criteria approaches. Fuzzy criteria entail using flexible alternatives instead of using a fixed set of query questions which only can be satisfied by a series of yes answers. One fuzzy criteria approach to find solutions when a search using all fixed criteria does not yield satisfactory answers is to eliminate some of the criteria, or base decisions on a count of how many factor measures meet the set criteria thresholds. Another fuzzy criteria option is to adjust fuzzy criteria thresholds up or down to qualify more or less component factors to satisfy the criterion.

A second search solution for multiple criteria searches can be to develop an index of weighted factors.

A third solution could be to utilize special information regarding the topic already developed by others with access to different information. Examples could be the use of packaged data solutions such as those provided by ESRI Acorn or Community Tapestry data. ESRI Business Information Solutions (2004)

Often complicated searches may be best undertaken using a mixture of techniques. For example simple yes-no queries using a database or spreadsheet program could be complemented by overlay screening queries in a GIS so that the spatial locations of areas satisfying queries could be determined. The case study described in this paper utilizes all of the techniques outlined in table 2.

Table 2: Techniques for Organizing Multiple Criteria Searches

- Simple yes-no query/overlay screening
- Fuzzy criteria adjustment to use counts of criteria meeting critical thresholds,
- Fuzzy criteria incremental adjustments to criteria thresholds
- Weighted indices
- Use of specialized additional information provided by other sources

2 OUR PROBLEM: FINDING NEIGHBORHOODS WITH "URBAN HIP" PROFESSIONAL CHARACTERISTICS IN THE SAN FRANCISCO BAY AREA OF CALIFORNIA

2.1 Urban Hip Professional Neighborhoods

Currently there is considerable excitement regarding a revival of interest in urban neighborhoods, particularly neighborhoods that have diverse, interesting mixes of people, developments and activities and good public transportation. Groups such as the Congress for the New Urbanism (2000) have promoted building liveable communities by fostering transit-oriented development (TOD). Robert Cervero, in **The Transit Metropolis – A Global Inquiry** (1998), reveals how density, diversity, and design are the three dimensions for the emerging "sustainable transit metropolises of tomorrow." Writings by Calthorpe (1993), Katz (1994), and others have identified places with at least moderate population and housing densities, and a variety of people with interests in social and political justice and alternative lifestyles as being what we call "urban hip" neighborhoods. Richard Florida (2002) has sparked considerable interest in urban hip professional neighborhoods with his provocative writings regarding the "Creative Class" and the relationships between economic growth and diverse, integrated, tolerant, live-work-learn-play communities. Florida (2005).

Urban hip neighborhoods are a subjective concept which can take a wide variety of forms and characteristics. For our case study example, we will look for urban hip neighborhoods in the San Francisco Bay region, with special attention to the needs of high tech professionals. An overview of the type of neighborhoods we are looking for has been provided by the **Los Angeles Times**. Under the headline "*The Newest Hip Suburb: Downtown - Life in the city core is luring many, who cite transit ease, social contacts, other benefits*" the **Times** went on: "Planners, developers, city officials and even transportation authorities say that maturing baby boomers and young people are migrating to downtown neighborhoods, where ...."townies" find a more energetic street scene and better access to transportation, jobs, and social outlets." Ailworth (2004)

Our case study seeks to find urban hip neighborhoods suitable for persons working in creative high technology jobs, with special attention to young singles and maturing empty-nester baby boomers. There is not space in this short paper to describe the logic behind the selection of the 12 measures outlined in Table 3, but it should be noted that all but two are readily available census data.

2.2 Attempted Use of Simple Yes-No Query/Overlay Technique to Find Tracts Meeting All Initial Criteria Thresholds

Typical solutions to find locations satisfying multiple criteria entail using database queries and GIS overlay maps. This can be called the "Simple Yes-No Query/Overlay Technique." Our initial attempt to locate urban hip neighborhoods in the San Francisco bay region was to start very broadly and construct census tract measures for each of 12 factors. (We assume that census tracts, or groups of tracts, represent neighborhoods.) For each factor we selected an initial threshold level above which the neighborhood would be considered hip. In most cases our initial assumption was to look for areas that ranked in the top 1/6 of the region and to use z-score transformations of the data (Z-scores represent the number of standard deviations each observation is from the mean of all observations.) The “Just Pass Threshold” data in table 4a and figure 1 present the values used to query all 1406 census tracts in the nine county San Francisco Bay region. Unfortunately the simple yes-no query/overlay test result was that no census tract passed even 11 of the 12 threshold tests. Perhaps our search may be a difficult one, comparable to the proverbial search for a needle in a haystack.

2.3 Fuzzy Criteria Query/Overlay Modifications to Find Tracts Nearly Meeting All Initial Criteria Thresholds

Fuzzy criteria standards are measures lacking exact certitude. As urban hip professional neighborhoods are very diverse cosmopolitan places, a wide variety of different mixes of people and developments may qualify a place as being hip. Our first fuzzy criterion was to drop the restriction that all neighborhoods must meet all criteria. Instead we would simply count the number of “yes” passes for each census tract meeting the initial urban hip criteria and call this sum the “meet initial criteria count” (IC). It is a fuzzy alternative to the failed strict yes-no query.

The first maps made for the entire nine county San Francisco region (not shown in this paper) revealed that much of the area away from the bay, especially the four north bay counties, has relatively low population and housing densities, and lacks green public transportation (public transit, walking, or bicycling). These circumstances doom those areas with poor urban hip prospects. Because of the meager urban hip professional prospects indicated by our maps for the north bay and outlying areas, our attention is concentrated on the portions of the five south bay counties bordering San Francisco Bay.

The red and tan colors in figure 1, following Tufte (1990) and Brewer (2004), show in detail the spatial patterns of factors meeting the initial urban hip criteria in the San Francisco Bay vicinity. Note the concentration of urban hip characteristics along the CalTrain commuter railroad line running from San Francisco to San Jose. The map of initial criteria test passed counts, top left of figure 1, shows a heavy concentration of areas meeting the urban hip criteria in San Francisco and a lesser but still significant cluster in the north west corner of Santa Clara County.

Placeholder for Table 3

Placeholder for Table 4

Placeholder for Figure 1

2.4 A Second Fuzzy Criteria Modification – Lower Threshold Criteria by an Increment

Figure 1 shows how well each Urban Hip criteria was satisfied, with areas identified as just passing and just failing the original threshold criteria by no more than the specified increment.

As our problem with the original Yes-No Simple Criteria test was that too few census tracts were passing the initial threshold criteria tests; we used Easy Urban Hip Criteria which enabled census tracts to qualify if they were no more than an increment below the initial criteria threshold. Under these conditions all of the “Just Fail” tracts on the Urban Hip Test Status maps would qualify. Table 4 identifies the increments by which threshold criteria could be changed to allow more or fewer areas to qualify as urban hip. For most measures the fuzzy increment was set at 0.5 standard deviations, which for normally distributed data would move about 8% of the areas into or out of qualification, depending on if the thresholds, were being raised or lowered. The light blue areas on the Figure 1 Urban Hip Initial Criteria Test Results maps indicate tracts which just fail under the initial threshold criteria, but would just pass if the threshold were lowered by the increment. We call the lowering of the threshold for qualifying as urban hip by a small increment the “Easy Criteria Test” (EC).

The initial and easy counts of threshold criteria met represent alternative ways of using fuzzy standards to select areas using multiple criteria. The initial counts may be narrow, but deep measures, as each qualifying standard is at least moderately strong. The easy counts may be broad, but shallow, as some qualifying measures may not be very strong. As there are advantages to both deep and broad measures, we chose to use both the initial and easy criteria counts as determinants of urban hip areas.

One problem with simple yes-no criteria counts is that they do not reflect the degree to which the criteria are being meet. For example it is possible for two areas qualifying as having green transportation to differ by a factor of two or three for the proportion of public transit users, walkers, and bicyclists, yet both areas score equally when using met criteria counts as the measure. This shortcoming can be overcome by use of a weighted index that can take account of the relative strengths of the factors. Our next step was to construct such an index.

Calculation of Weighted Z-score Summary Index Measure

Table 5 documents the production of a weighted z-score summary index measure. Z-scores are informative ratio levels of measurement which, when calculated for neighborhood areas, can inform us how different these areas are from the average value for all areas in the region. Z-scores for large data sets approximating a normal distribution have the added advantage of implying percentage shares or proportions. For example, 2/3 of all z-scores for normally distributed data will be within plus or minus one z-score (standard deviation) and 95 percent within two z-scores.

Table 5 outlines, following approaches advocated, by Babbie (2004), how we weighted the index to take advantage of under represented factors, such as giving a weight of 3 to college education and provided maximum caps on many z-score values to prevent extreme values from overwhelming the index.

2.3 Refining Census Tract UHIP Rankings to Account for Cluster Aggregation Economies

Clusters of urban hip tracts are likely to be more desirable than tracts standing alone with comparable socio-economic-demographic characteristics. These clusters may support larger and more varied services and may enable the cluster area to cross minimum market area thresholds which will enable some urban hip businesses to prosper where otherwise they may not have sufficient customers in more isolated areas. Figure 2 describes how proximity bonuses were assigned for census tracts within _ mile of tracts with top classifications on the three aggregate UHIP measures we prepared. The calculation of the bonuses was based on the presence of tracts within a _ mile buffer surrounding all tracts which ranked high on all three aggregate measures. The bonus calculation also utilized counts of tracts with high scores on the initial and easy criteria counts, as well as the z-score index.

Our UHIP aggregate census tract ratings were adjusted by adding the proximity bonuses to the original counts of initial and incrementally adjusted threshold criteria met and the weighted z-score index.

Place holder for Table 5

Places for Figure 2

Placeholder for Figure 3

3 FINDING GOOD URBAN HIP CENSUS TRACT PROSPECTS USING FOUR AGGREGATE MEASURES

3.1 Creating a final composite measure by combining 3 aggregate measures

The upper or left three maps in figure 3 present the initial results of our search for Urban Hip neighborhoods in the San Francisco Bay region. The bottom central ”Final Ranks” map is the average of the ranks found for the data used to produce counts of the passes on the initial and easy criteria tests and for the z-score index, with all measures including proximity effect adjustments.

“Excellent” UHIP prospects on the criteria counts required adjusted scores of greater than 8.01 for passes using the initial criteria (IC) and 10.01 for passes using the easy criteria (EC). The “very good” UHIP prospects shown in figure 3 for the criteria counts require adjusted count scores 2 points lower.

Only 17 tracts had excellent prospects according to the IC tests, while 31 tracts had excellent prospects using the EC count measure. Very good prospects were foreseen for 37 tracts using the IC, while 39 tracts were given similar prospects by the EC count from among the 1,406 tracts in the greater San Francisco region. (Remember that the easy counts, while having factor thresholds usually lower by half a standard deviation, required two more pass counts than the harder IC to achieve similar prospects.)

The z-score index, bottom left in figure 3, shows similar patterns for excellent and very good prospects to the patterns shown above it by the criteria counts. Figure 3 clearly identifies the strongest clusters of urban hip professional prospects in the region as being in the City of San Francisco. However the northwest corner of Santa Clara County has a significant cluster of good to excellent urban hip prospects of its own. The bottom right map of figure 3 zooms in to detail the final ranks for what we call the Silicon Valley UHIP Cluster. This loose cluster lies along the CalTrain commuter railroad line between downtown Palo Alto (census tract 5113) and downtown Mountain View (census tract 5096). A solid brown patch of very good to excellent urban hip prospects is evident between these two downtowns on the EC Counts map, but the pattern is weaker on the IC Counts map. This suggests a broad array of urban hip features but perhaps not heavy relative concentrations of many measures.

Comparison of the maps across the top of figure 4 reveals that the multiple family housing, high density housing, and green transportation urban hip measures closely correspond to the ESRI Tapestry urban hip type classifications; this is especially true for Laptops and Lattes. Our service and retail jobs per square mile patterns also compare well with the Tapestry data. Our in-mover diversity measures generally are quite strong in the areas identified with promising urban hip prospects, but are strong elsewhere as well, indicating that in-movers are likely to be a necessary but not sufficient condition to identify urban hip areas.

3.2 Using Special Additional Information

Figure 4 focus on the areas passing the initial criteria tests by showing the easily pass and just pass conditions in bold colors and showing all the areas which fail to pass in a quiet dull tone. This figure also maps ESRI's Community Tapestry data classifications, ESRI (2004), likely to be compatible with urban hip status. ESRI's "Laptops and Lattes" classification was expected to most closely correspond to our urban hip professional classification and in fact all 25 of our top ranked UHIP tracts in the region had this classification. A comparison of the maps across the top of figure 4 reveals that the multiple family housing, high density housing, and green transportation urban hip professional measures closely correspond to the ESRI Tapestry urban hip professional type classifications; this is especially true for Laptops and Lattes. Our service and retail jobs per square mile patterns also compare well with the Tapestry data. Our in-mover diversity measures are strong in the areas identified with promising urban hip professional prospects, but are strong elsewhere as well, indicating that in-movers are likely to be a necessary but not sufficient condition to identify urban hip professional areas.

For our analysis David Roemer showed how photos linked to a map can provide a qualitative dimension to envisioning a place. He downloaded the City of Mountain View's general plan land use map for the downtown area from the web and surrounded the map with photos documenting the interesting variety of places and activities in the area.

4 CONCLUSION

Spatial queries using multiple fuzzy criteria may be easy or difficult depending upon the conditions and the data. Using spreadsheet and database queries, complemented by multiple GIS maps we have sought to find neighborhoods with multiple special characteristics. Modern technology has made the problem of finding a needle in a haystack easier to solve. In GIS we now have a converter belt to quickly spread the haystack out into a single layer so we can see the bright and shiny needle.

Placeholder for Figure 4

We also have spreadsheets with logic statements and database queries that work like magnets to grab the needle without us even seeing it. This however leaves us with the more difficult searches that can be compared to finding a needle in a haystack which also has a handful of pins. While technology can help in that search, there isn't a complete, easy technological solution, yet. The search is still likely to require careful individual examination of each selected object with a distinctive eye to see if it is indeed the needle and not a pin. In seeking urban hip professional neighborhoods, we may use our modern research techniques to limit the search to selected likely prospects, (i.e. pins and needles) but an effective ultimate choice is nevertheless likely to require direct personal observation.

We have demonstrated that using fuzzy criteria searches can help find good prospects when initial simple yes-no queries fail to yield results. Fuzzy criteria can include using counts of criteria test satisfied, rather than requiring full passing of criteria. Other fuzzy criteria techniques may entail lowering criteria pass thresholds but requiring more criteria to be satisfied. When time and resources permit, developing weighted indices may more effectively measure intensity of factors, a detail that simple counts of success may overlook. In addition spatial searches should consider cluster effects, something that can be readily done using buffers to develop proximity or agglomeration cluster measures.

To sum up, technology can find that needle in a haystack, but if the haystack also has some pins of comparable size and color then individual examination of objects after spatial filtering is likely to be a final part of the search. As for the more difficult problem comparable to finding a needle in a barrel of pins, that is a challenge we can look forward to solving using Longhorn and the next generation of data mining tools.

ACKNOWLEDGEMENTS

The decade long research into envisioning neighborhoods has been assisted by numerous persons including students and faculty at San Jose State University. Grants from the Mineta Transportation Institute at San Jose State University and released time grants and leaves from San Jose State University have been helpful, as has support from a National Science Foundation grant (0228878) overseen by Richard LeGates at San Francisco State University. ESRI and ESRI Press have provided support over the years for the Envisioning Neighborhoods efforts. For this paper vital GIS help was provided by University of California, Davis Ph.D. candidate Xuyang Zhang. David Roemer played a major editorial role for this paper. The figures used the 11-class diverging red-blue and 11-class diverging brown-blue-green color schemes from Cynthia Brewer's ColorBrewer website referenced below.

Note: This paper is a revised version of a similar paper prepared for the Computers in Planning and Urban Management (CUPUM) conference to be held in London, England in June, 2005.

REFERENCES

a) Books and Books chapters

Babbie, Earl (2004) **The Practice of Social Research**. Wadsworth Publishing Company, Belmont, CA.

Calthorpe, Peter (1993) **The Next American Metropolis – Ecology, Community, and the American Dream**. Princeton Architectural Press, New York.

Cervero, Robert (1998) **The Transit Metropolis – a Global Inquiry**. Island Press, Washington, D.C.

Congress for the New Urbanism (2000) **Charter of the New Urbanism**. McGraw-Hill, New York.

Florida, Richard (2002) **The Rise of the Creative Class – And How It's Transforming Work, Leisure, Community and Everyday Life**. Basic Books, New York.

Florida, Richard (2005) **The Flight of the Creative Class**. Harper Business, New York.

Katz, Peter (1994) **The New Urbanism – Toward an Architecture of Community**. McGraw-Hill, New York.

Tufte, Edward (1990) **Envisioning Information.** Graphics Press, Cheshire, CN

c) Papers presented to conferences

Bossard, Earl and David Roemer (2004) Envisioning Neighborhoods: GIS Information Design to Facilitate Understanding Places, **ESRI 2004 International User's Conference Proceedings** at URL:http://gis.esri.com/library/userconf/proc04/docs/pap1525.pdf

d) Other documents

Association of Bay Area Governments (2002) Projections 2003 by Census Tract. ABAG Publications, Oakland, CA.

Brewer, Cynthia. (2005) **ColorBrewer – A web tool for selecting colors for maps. http://www.ColorBrwer.org, accessed April 6, 2005.**

Bossard, Earl (2002) **Envisioning Neighborhoods with Transit-Oriented Development Potential**. Mineta Transportation Institute Report 01-15. San Jose: Mineta Transportation Institute, College of Business, San Jose State University. URL: http://www.sjsu.edu/depts/urbplan/up0d131.htm

City of Mountain View General Plan Land Use Map. http://www.mtnview.ca.us/citydepts/cd/cp/pdf/CMV_2005_General_Plan_Land_Use_Map.pdf

ESRI Business Information Solutions (2004) Introduction to Community Tapestry Segmentation. ESRI Business Information Solutions, Redlands, CA.

U.S. Census Bureau (2002) 2000 Census of Population and Housing, Summary File 3: Technical Documentation. Washington D.C.

AUTHOR INFORMATION:

Earl G. BOSSARD
Professor
Urban and Regional Planning Department
San Jose State University
San Jose, CA 95192-0185
Tel: 530 758 1602
Tel: 408 924 5860
E-mail: bossard3@pacbell.net
Website: www2.sjsu.edu/urbanplanning/SJSU/people/EBossard