THE POWER OF SYMBOLOGY IN THE GIS WORLD

Michael E. Gentles
Senior Consultant
Synercom Technology Inc.
Houston, Texas 77042

ABSTRACT

This paper will examine and explain some unique ways for analyzing data with the use of symbols. Symbolic maps are an effective way to display spatial information. Symbols can show a number of variables with the same display. Sixty to seventy percent of the data within an organization is geographic related and has a geographic identifier. Types of data that can be displayed using creative symbology are land-information/parcel data, socio-economic data, load studies using meters and transformers, and incident reports for crime, such as the locations where cars were stolen and recovered and violent crime has ocurred.

When analyzing data with a geographic spatial placement and location, the power of the symbol is much greater than the standard 2-dimensional or 3-dimensional perspectives of thematic analysis. This paper will explain the uses of symbol shape, size, and color to assist in the analysis of data.

INTRODUCTION

The use of symbology with automated mapping and facilities management in the areas of data analysis, data validation, and quality control provides an excellent and powerful tool in which to communicate to the map reader additional underlying phenomena about a spatial feature. Maps can be very simple showing a two dimensional difference or very complex, showing many variables within a symbol. As the maps become more complex they must be able to show the quantitative geographical phenomena as spatial attributes.

The examples and concepts used throughout this paper will concentrate on maps that range from a scale of 1:500 to 1:5000. The examples and data bases will come from municipalities using multipurpose cadastre, public works departments, and utility companies with whom the author has worked with in the past.

BACKGROUND

With the rapid expansion of computerized mapping using attribute information attached to the graphic symbology, the means for improved data analysis and data validation can be accomplished with the use of symbology. Large tabular listings used in the past for data analysis and data validation now can be replaced or supplemented by multi-symbolized maps to assist the analyst or quality control personnel in checking and analyzing data.

But where does this attribute data come from? The attribute data attached to the graphic symbology comes from a number of different sources, including existing hardcopy, manually drawn maps, or bulk loaded from computerized files. Once this attribute data is attached to the symbology there are still some very important questions that need to be answered. How valid was that data from the original sources? What checks were used to analyze the quality of the data prior to being attached to the data base? Is the data spatially correct?

Through the use of computerized mapping systems, the user has a choice of displaying the data in a number of ways when boolean criteria is applied against the attribute data.

SYMBOLOGY

Symbology on maps are displayed in three classes;

Point	Representing a single geographic location, such as the visual centroid of a lot, a location of a device such as a electrical transformer, or water valve
Line	Representing a linear feature such as an electrical primary conductor, a water line, or a shoreline
Area	Representing a geographic area (polygon), such as a census tract, the area of a lot, or other thematic presentation

There are two different types of symbology, replicative and abstract. Replicative symbols are those that are designed to look like their real world counterparts; they are only used to represent tangible objects. Abstract symbols generally take the form of geometric shapes such as circles, squares, and triangles. (2 p.20) Coast lines, trees, houses, and cars are examples. Base map symbols are replicative in nature, whereas thematic-overlay symbols may be either replicative or abstract.

The visual representation of a symbol when built against the attribute data, can tell the user many different things about the symbol. This is done by using the shape, the size, the color or any combination of these representation. The shape of a symbol has already been addressed. The size of symbol can be varied as to a given set of ranges. The lower values of a range would be shown with a smaller sized symbol, while a larger sized symbol would indicate a larger value. Color can easily be used to represent different features or be used to show different ranges.

EXAMPLES

A symbol within a lot or parcel of land can represent a number of different things when the symbol is determined by one or more of the associated attributes of that parcel centroid. Table I. illustrates how the number of bedrooms for a parcel could be shown.

TABLE I

Number of bedrooms	Shape	Color	Size
2 or fewer	plus	green	one half normal size
3	triangle	blue	normal size
4	square	red	one and a half times normal size
5 or more	star	purple	twice normal size

Using either the shape, the color, or the size of a symbol, a user analyzing a map portraying the number of bedrooms would easily be able to distinguish by the symbology which parcels had two bedrooms, three bedrooms, four bedrooms or five or more bedrooms. When spatially analyzing a large geographic area an anomaly will stand out. If a neighborhood had homes with only two and three bedroom homes, a five-bedroom symbol could warrant additional research as to the validity of the data. There may be other underlying data that would validate the number correct.

In Table II, parcel number four, indicates eight bedrooms, but further examination of that record reveals the parcel was zoned for a four plex, indicateing eight bedrooms is feasible. If parcel number four had been zoned as an R-1 single-family residence, then the number of bedrooms could require verification.

TABLE II

Parcel Number	Zoning Code	Land Value	Improvment Value	Square Feet	# of Bedrooms
1	R-1	$22,000	$45,000	1800	3
2	R-1	$19,800	$48,000	1467	2
3	R-1	$10,001	$55,000	1782	4
4	R-4	$35,000	$75,000	2412	8
5	R-1	$21,000	$42,000	2162	3
6	R-1	$29,900	$44,000	1944	3
7	R-1	$22,000	$51,000	1827	4
8	C	$60,100	$83,000	3775	-

When analyzing the maps, more than one dimension is needed to analyze the underlying reasons for an apparent anomaly. As in the above example, if zoning was to be shown by one method, e.q. color, number of bedrooms by the shape of the symbol, and the size of the structure by the scaled size of the symbol, it would be simple to determine if the data appeared correctly.

With the use of computer-generated maps, a single symbol within a lot or parcel could portray a number of different variables, such as:

- range of improvement value
- the square footage of the residence
- the number of bedrooms

The range of improvement value could be shown by using different colors to indicate the various ranges. A green symbol could represent an improvement value less than $40,000.00, a cyan-colored symbol would represent a range of $40,000.00 to $50,000.00, a magenta-colored symbol would represent an improvement value of $50,000.00 to $60,000.00, and a red-colored symbol would indicate that the improvement value was greater than $60,000.00. The square footage of the residence would be shown by the scaled size of the symbol, a smaller size symbol indicates a low square footage, while a bigger size symbol indicates a larger square footage residence. The number of bedrooms would be shown by the shape of the symbol. Two bedrooms and less would be indicated by a circle, a three-bedroom residence would be shown with a triangle, a four-bedroom residence with a square, and a residences with five or more bedrooms would be shown with a star. This same symbol symbology could be made even more complex by combining a second symbol on top of the first using different colors, different scales, and different shapes to add three other variables to the one symbol location.

Using the choropleth/thematic mapping method to indicate land values will not show any large discrepancies. As an example, creating a land value choropleth map (Figure 1), from the attribute data from Table 2, and using the ranges in Figure 1, $0.00 to $9999.99 is shown in the lightest texture, thru $50,000.00 and above indicated by the darkest texture. Two adjacent parcels appearing to be the same size, show in two different textures. The difference in the land value between these two parcels could be $1.00 apart or $19.999.00 apart. The reviewer or users need some way to evaluate the validity of the data. If the users had used symbology with the size of the symbol to indicate the land value and the symbol shape to indicate zoning, as in Figure 2, any large discrepancies could be easily distinguishable and possible conflicts resolved.

Figure 1

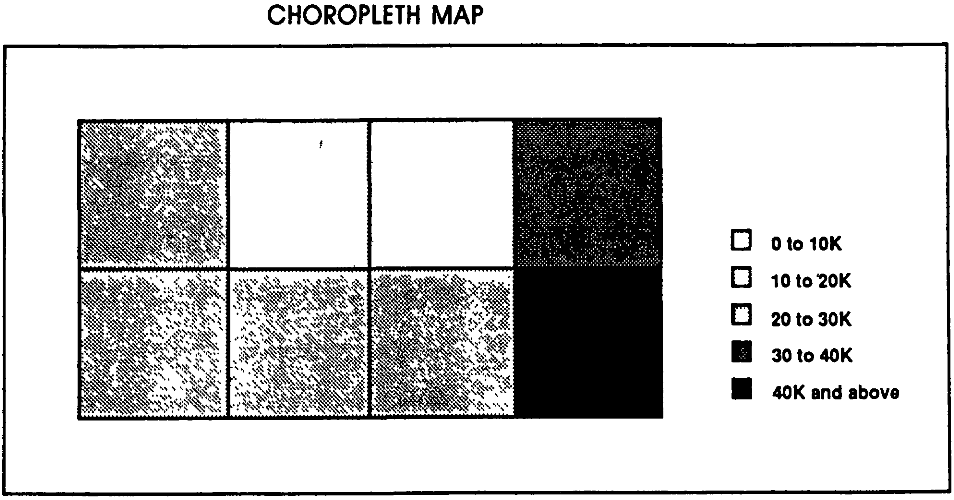
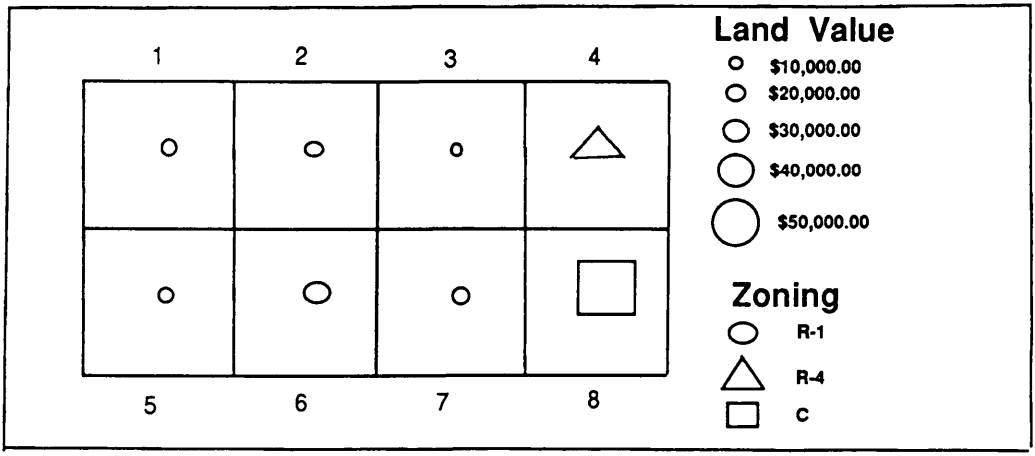


Figure 2



Another example of the use of symbols for data validation is with the use of linear maps. Linear maps can also have complex symbology to indicate two or three different types of symbology. Electrical distribution will be used in the following example and will include color for the electrical phase or phases, the thickness of the line will indicate the KVA of the line and the symbology within the line will indicate the material of the line.

In working with electrical distribution networks, there are three phases for the primary conductor, A, B, and C. Using the three primary colors, red, yellow, and blue one can create line colors that will visually point out any phase errors very quickly. If all A phase lines are done in red, all B phase lines are done in Yellow, and all C phase lines are done in Blue, then the following colors are created

- ABC = white
- AB = orange
- AC = violet
- BC = green
- A = red
- B = yellow
- C = blue

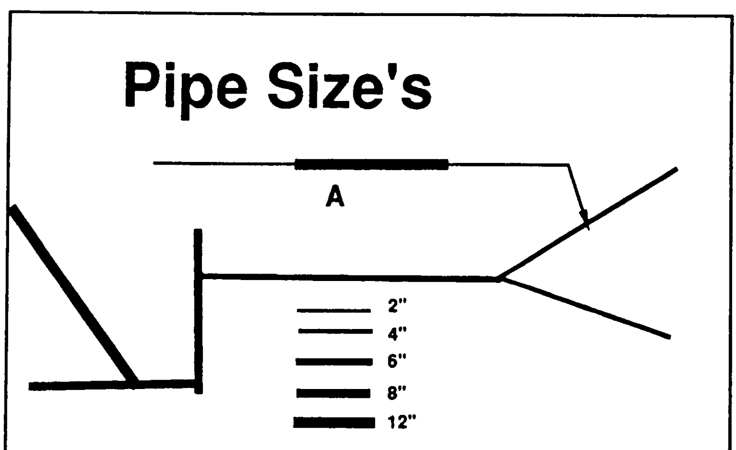
With this color scheme set up, the user can now visually see that the phase of the electric lines and the associated electrical devices have the proper phase conductivity by working with the following rules;

- A white line can have any combination of device colors attached to the line or branching from the line.
- An orange line is fed by a white line and can only have a red, orange, or yellow device attached or branching from the line.
- A violet line is fed by a white line and can only have a red, blue, or violet device attached or branching from the line.
- A green line is fed by a white line and can only have a blue, green, or yellow device attached or branching from the line.
- A red line can only be fed by a white, orange, or violet line and can only have a red device attached or branching from the line.
- A yellow line can only be fed by a white, orange, or green line and can only have a green device attached or branching from the line.
- A blue line can only be fed by a white, green, or violet line and can only have a blue device attached or branching from the line.

The type of primary KVA line can be symbolized by a line pattern indicating whether it is a low, medium or high-rated KVA type of primary. A 2-KVA line would be indicated by the line being one pen width wide. The 7-KVA lines would be portrayed as two pen widths wide and the 14-KVA primary lines would be portrayed as lines being three pen widths wide. At the locations where the line width changes, indicating a change in the KVA, there must be a step down/up device at the location. If there is not a step down/up device at this location there is an apparent error either in the data or in the conversion effort.

Another example of using linear maps to check out the data attribute of size for water and/or gas lines is to plot each different size pipe in a different color or line width. By plotting the different size water lines with different shading, line widths, or colors, sizes that appear to be incorrect such as an 8-inch water line connected on each end to a 2-inch line would stand out. This is indicated in Figure 3 next to the 'A', and that piece of 8-inch pipe would be marked for further investigation. As mentioned above, with the electrical devices being plotted in different colors to ensure correct phasing, all devices attached to the pipe lines such as valves, regulators etc. would have the size of the device plotted in the same color to help ensure the validity of the data attached to the pipe lines.

Figure 3



There are limitless other unique types of plots using different types of symbology that can be produced. Recently I had a chance to review some crime incident plots produced for a municipality that indicated the type of crime. What was unique about these plots was the symbology used. Armed robbery was shown with a symbol the shape of a gun, a stabbing was shown with a knife, stolen cars symbol was a car. To assist in the analysis of these crime, the time of day could have been indicated by the color, leting the watch commander know if the crime was being committed on his shift.

Another municipality uses the size of the symbols to indicate water pressure for all of the fire hydrants. This allows the fire department when responding to a call to use pressure as well as proximity to the fire, to determine which hydrant to use for the fire.

A municipality on the west coast has all the Class A petroleum pipelines, those pipelines that are highly flammable, that pass through the city displayed in red. The materials carryied through the line is plotted by using different line patterns.

CONCLUSION

What we have attempted here is to indicate how symbology can be used for data analysis and data validation. The options in working with multiple symbology using shape, size, and color, is almost limitless and the author encourges the readers to experiment with their own data in producing different types of maps.

REFERENCES

(1.) Auto Carto IV Vol II p158 "Interactive Computer Mapping Applications to Criminal Justice Planning in Three Virginia Cities"
Leonard J. Simutis, Assistant Dean
Todd A. Scott, Instructor
Division of Environmental and Urban Services
Virginia Tech, Blacksburg Virginia 24061

(2.) Principles of Thematic Map Design. Borden D. Dent
Copyright 1985 by Addison Wesley Publishing Company, Inc.