Tutorial for Topological Representation of Cities Based on Wholeness

Zheng Ren and Ding Ma
Faculty of Engineering and Sustainable Development, Division of GIScience
University of Gävle, 801 76 Gävle, Sweden
E-mail: renzheng1991@gmail.com

1. Introduction

This tutorial aims to provide an instruction about how to visualize the cities as a topology-oriented representation based on the view of wholeness. All cities within a country possess different degrees of wholeness (Jiang, 2015a), which can be characterized by a novel geographic representation. Firstly, this representation abstracts all cities as individual points and symbolizes those points into different hierarchical levels according to their sizes. The symbolization is conducted with the help of HT Mapping function in the Axwoman 6.3 (Jiang, 2015b), based on head/tail breaks – a new classification scheme and visualization tool for data with a heavy-tailed distribution (Jiang, 2013). Secondly, those points in different hierarchical levels are regarded as the input shapefile to create Thiessen polygons. After deriving those Thiessen polygons, we build up a complex network to calculate the degree of wholeness based on polygon – polygon relationships. The network follows the rules that small polygons point to their adjacent large polygons within the same hierarchical level and contained polygons point to their containing polygons across two consecutive hierarchical levels (Jiang, 2016). Finally, we can compute the degree of wholeness based on this network for individual cities and visualize the cities in a radial layout. This kind of representation is topology-oriented which is distinctive from the geometry-oriented representations. It concerns that taking all cities as a whole to uncover the underlying hierarchical structure of cities in the geographic space. Furthermore, it can be applied to other designs or patterns to assess their beauty of structures.

The reminder of the tutorial is structured as following parts: preparing the data and software that will be used in the tutorial, creating Thiessen polygons in ArcGIS, constructing directed graph based on the relationships of Thiessen polygons, calculating degree of wholeness in Matlab, and visualization of the hierarchical levels of cities in form of radial layout. Finally, you can find out that the degrees of wholeness possess both properties of differentiation and adaptation by this kind of representation. (Note: the symbol '>>' means next step of operation.)

2. Prerequisites

2.1 Data

The data you need including three shapefiles:
(1) The original point features to for creating the natural cities, for example, the street nodes.
The sample data provided is: **StreetNodes.shp**
(2) The natural cities in a polygon feature class.
The sample data provided is: **NaturalCities.shp**
(3) The boundary polygon of your process area.
The sample data provided is: **Boundary.shp**
Note: Here is a tutorial about how to extract street nodes and create natural cities. The ArcGIS models and the sample data are also provided: http://www.arcgis.com/home/item.html?id=47b1d6fdd1984a6fae916af389cdc57d

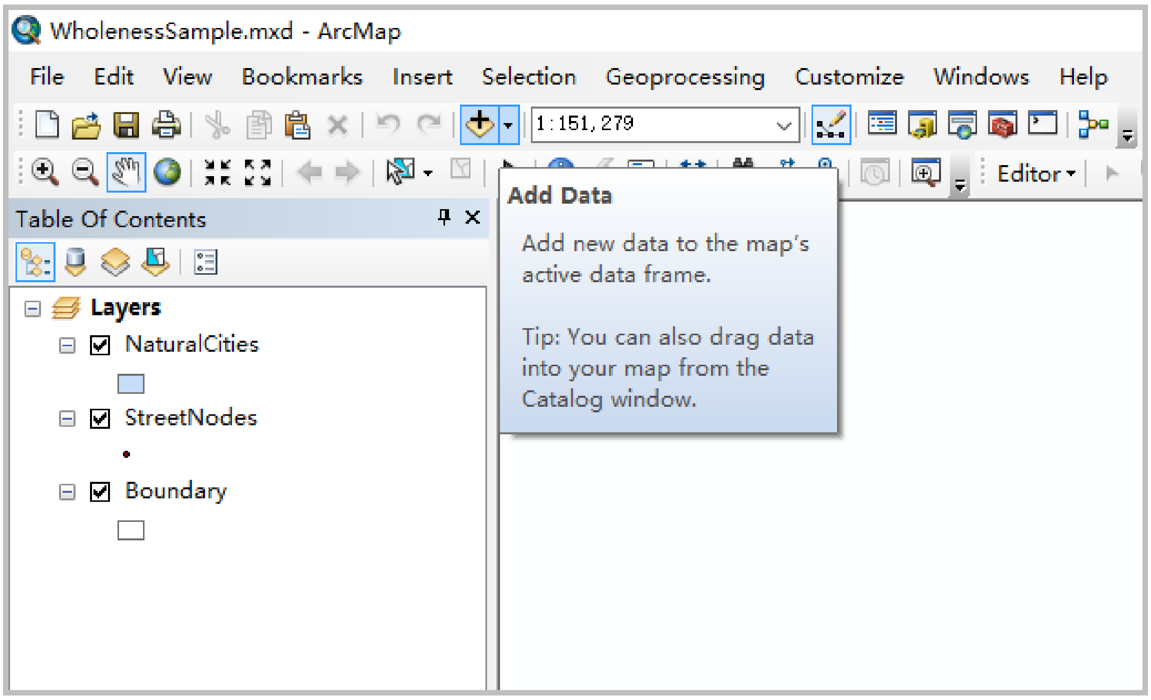
2.2 Software

There are five categories of software you will need:
(1) ArcGIS 10.2 with Axwoman 6.3 extension and ArcGIS SDK .NET 10.2 installed.
You can download the Axwoman here: http://fromto.hig.se/~bjg/axwoman/#tutorial
(2) Matlab 7 for calculating the degree of wholeness.
The degree of wholeness can be characterized by Pagerank (Jiang, 2015a), the code for the calculation of the Pagerank in Matlab 7 is provided by David Gleich (2006). You can download the Matlab code here: www.mathworks.com/matlabcentral/fileexchange/11613-pagerank
(3) A small program which is used to create the directed graph. You can download the program from this link: https://github.com/ZhengRen91/Topological_Representation_CreateGraph
(4) The Excel is used to sort the data from the text output from the constructing network program.
(5) A radial layout program for the visualization of directed graph
This program is developed by Ding Ma, and you can download the program from the following link: https://github.com/digmaa/Topological-Representation-Radial-Layout

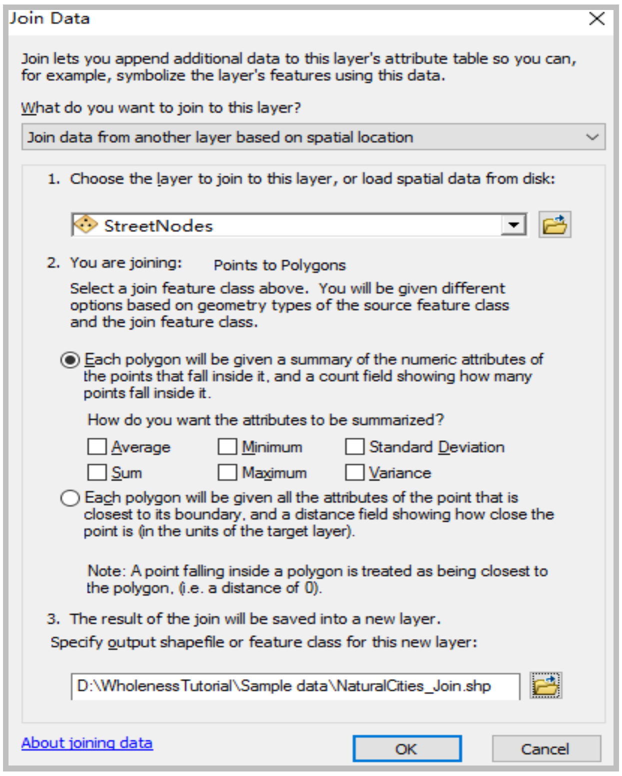
3. Creating Thiessen polygons in ArcGIS

Firstly, we need to create Thiessen polygons from different levels of cities, which are classified according to the number of points inside the natural cities based on head/tail breaks. Three sample shapefiles have been provided. Before you start, create a working folder for all processing data.

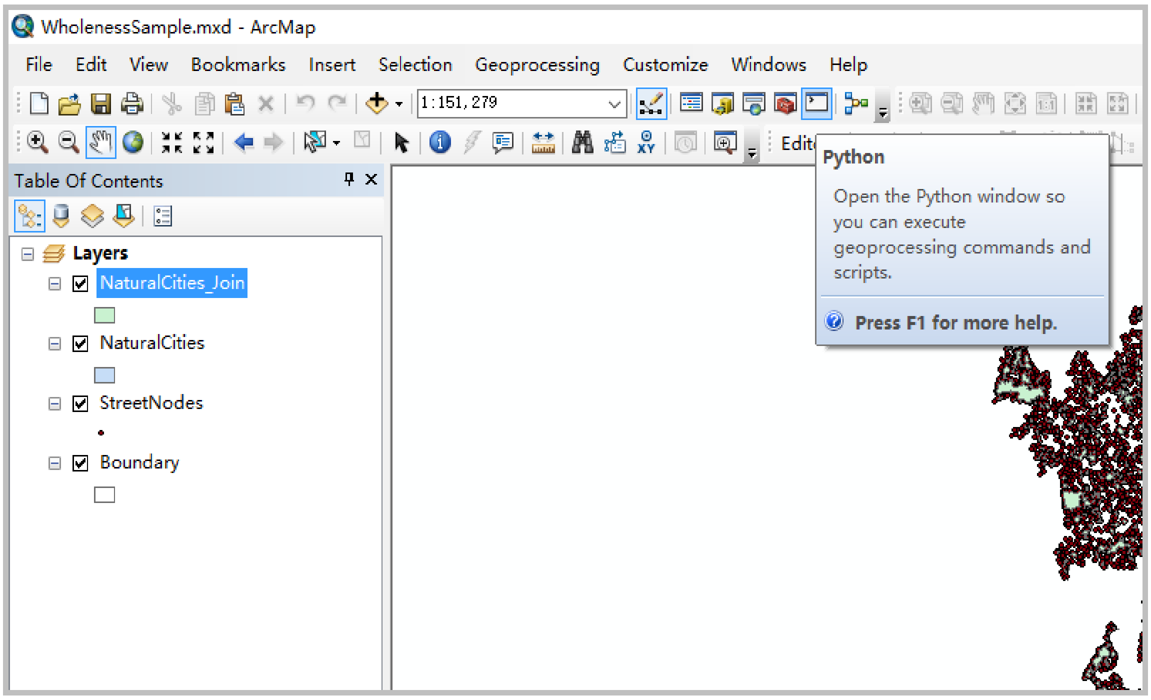
1. Open ArcMap >> Add three sample shapefiles from your folder to the table of content.



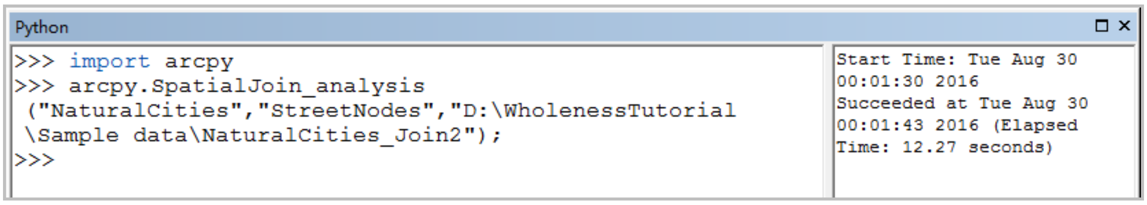
2. Get the number of street nodes in each natural city polygon. Right click NaturalCities layer >> Choose Joins and Relates >> Join... >> Open Join data window >> Choose 'Join data from another layer based on spatial location' >> Choose StreetNodes as another layer >> Specify the output shapefile as 'NaturalCities_Join.shp'.



This join process can be slow when the data is big. Alternatively, you can use Python window in the ArcGIS to speed up the join operation. Click Python window in the standard tool bar.

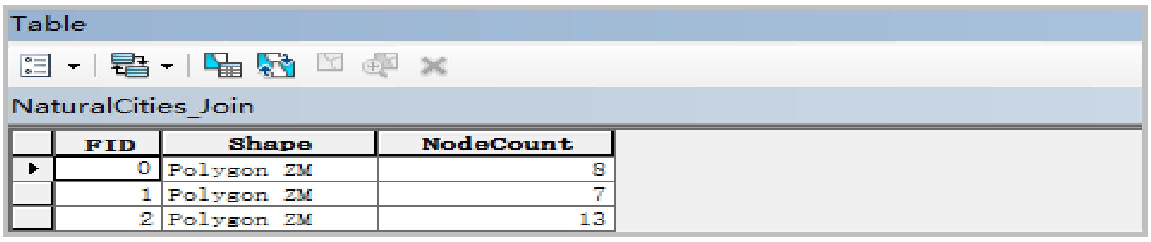


In the Python window, input the geoprocessing commands as shown in following figure >> Press enter to start.



3. We need to tidy up the attribute table to make it neat.

Right click NaturalCities_Join layer >> Open Attribute Table >> Add a new interger fied called NodeCount >> Right click new field >> Field Calculator >> Input NodeCount = Count_ >> Right click Count_ field >> Delete Field >> Right click Id field >> Delete Field >> OK. The Columns should look like as the following figure.

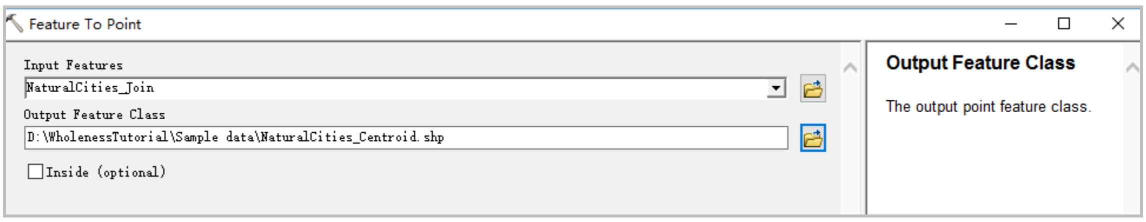


Table

NaturalCities_Join

FID	Shape	NodeCount
0	Polygon ZM	8
1	Polygon ZM	7
2	Polygon ZM	13

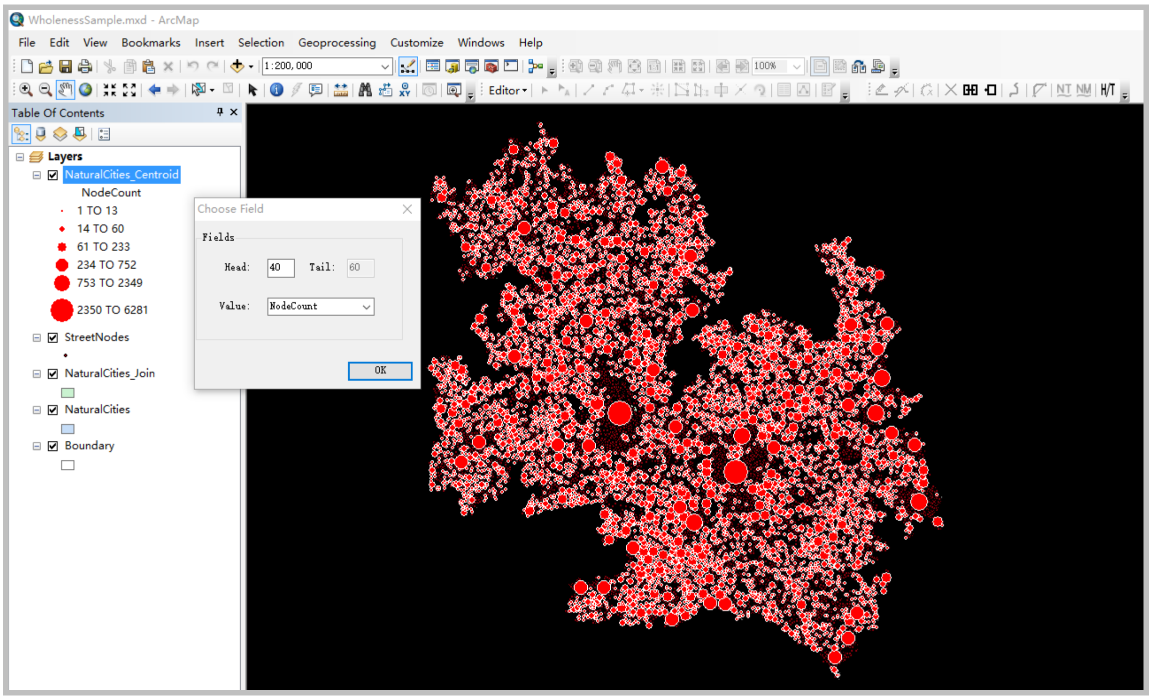
4. Now we abstract the natural cities polygons as centroid points. Open ArcToolBox >> Data Management Tool >> Features >> Feature to point. Specify the input features as NaturalCities_Join and the output feature class as NaturalCities_Centroid, leave the Inside option unchecked and click OK to start processing.



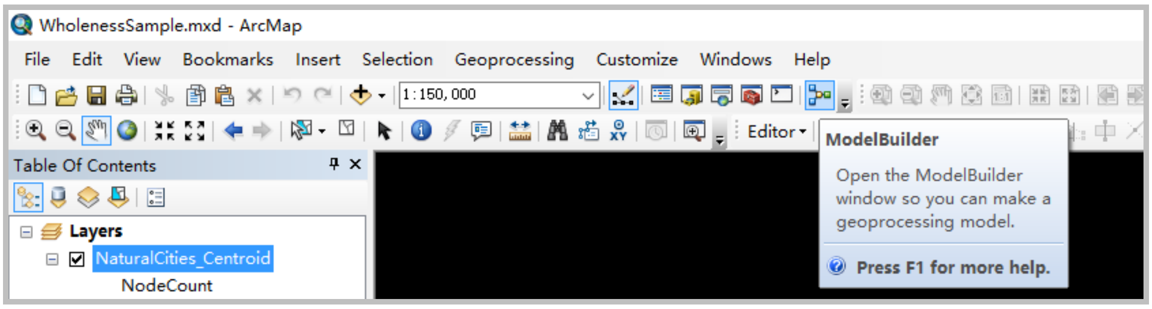
Note that the output shapefile should contain the ORIG_FID and NodeCount columns.

FID	Shape *	NodeCount	ORIG_FID
0	Point ZM	8	0
1	Point ZM	7	1
2	Point ZM	13	2
3	Point ZM	8	3
4	Point ZM	11	4
5	Point ZM	18	5

5. Use H/T mapping button in Axwoman to classify the centroids according to the numbers of street nodes. Click to select the NaturalCities_Centroid layer in the table of contents >> Click H/T button in the Axwoman toolbar >> Keep head and tail percentage as default and choose NodeCount as Value >> OK.

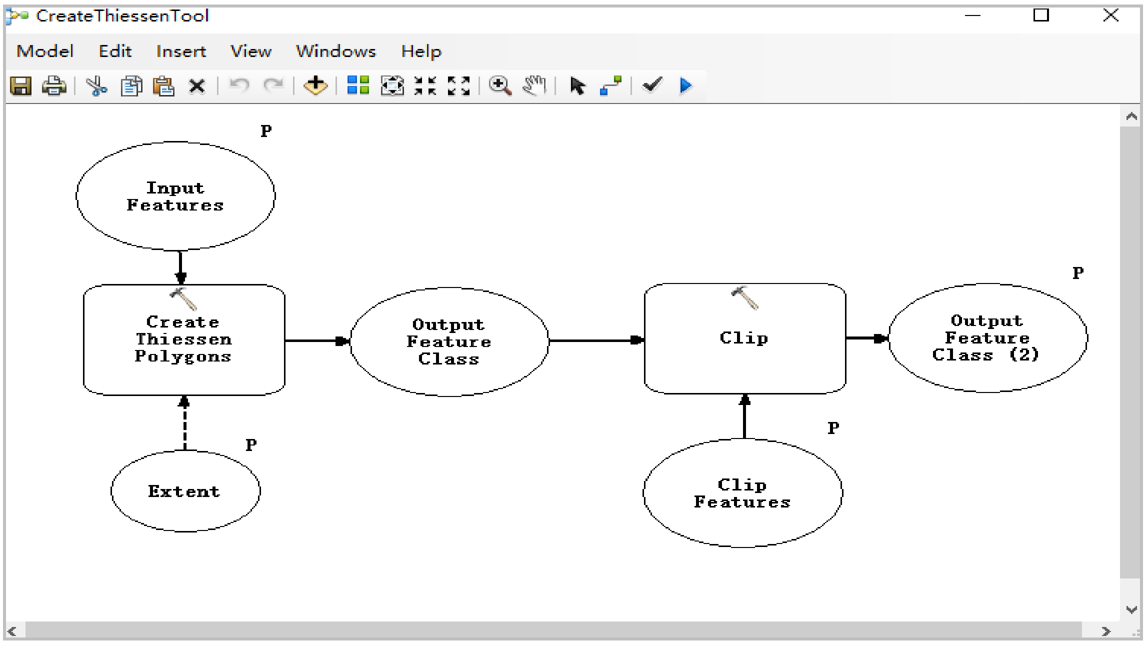


6. There are 6 hierarchical levels in this case. In order to speed up the process, we can create a simple model uising ArcGIS model builder. In the standard toolbar, click to open the model builder.

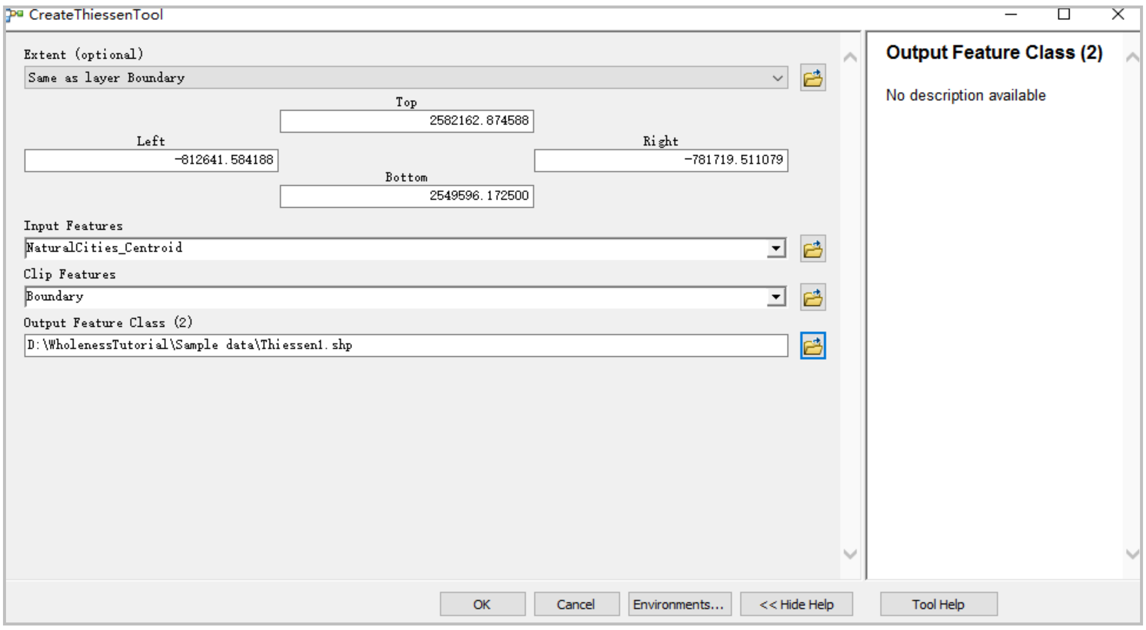


Open ArcToolbox >> Analysis Tools >> Proximity >> Drag and drop Create Thiessen Polygons in the model window >> Extract >> Drag and drop the Clip tool in the model window >> Use Connect tool in the model builder to link the Output Feature Class of Create Thiessen Polygons to the input of Clip. >> Right click Create Thiessen Polygons >> Make variable from enviroment >> Processing Extent >> Extent >> Right click Extent >> Tick Model Parameter, there will be a ‘P’ beside the Extent circle. >> Right click Create Thiessen Polygons >> Make variable from parameter >> Input Features >> Right click and tick Model Parameter >> Right click Clip >>

Make variable from parameter >> Clip Features >> Right click and tick Model Parameter >> Right click Output Feature Class of Clip and tick Model Parameter >> Double click the CreateThiessenPolygons and choose 'ALL' in the Output Field. Note that last step is very important for it maintains the NodeCount information in the output Thiessen polygons. The model is shown in the following figure, you can drag polygons to adjust their locations. The tool should be saved into a toolbox, you can create a toolbox in your folder and save this CreateThiessenTool in it.



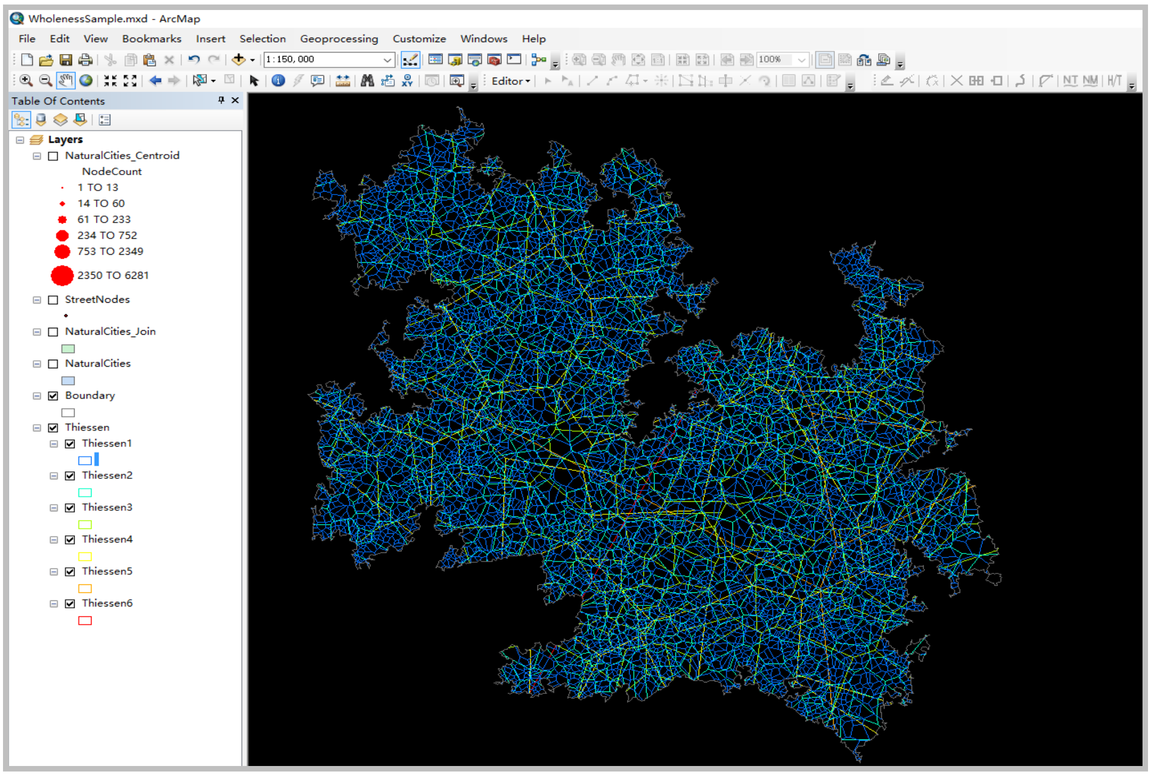
7. Now we select first level of centroid according to the H/T mapping result. Right click NaturalCities_Centroid layer >> Open attribute table >> From table option click select by attributes >> Type the condition: "NodeCount" >1 AND "NodeCount" <=13 in the condition box >> Click Apply. In this case 5990 out of 7084 centroids were selected. Open the CreateThiessenTool >> Specify the Boundary as Extent and Clip features >> Specify the selected NaturalCities_Centroid as Input >> Name the output as 'Thiessen1.shp' in your folder >> Click OK to start.



8. Right click NaturalCities_Centroid layer >> Data >> Export data >> Export selected data >> Name the output

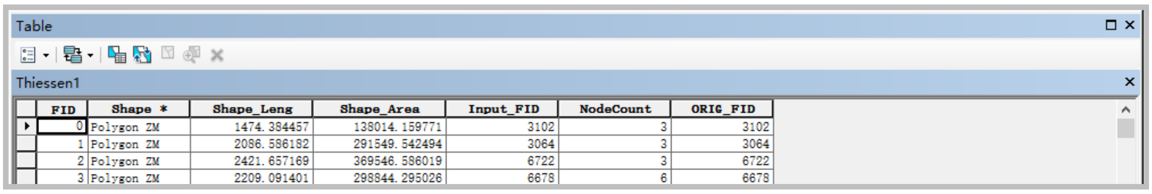
feature class as ‘Centroid1.shp’. This step is important, we need different levels of Centroids to construct graph.

9. Repeat step 7 and 8 until all levels of centroids are processed. Note the relational symbols of the selecting condition, for example, the condition of second level should be "NodeCount" >13 AND "NodeCount" <=60. Finally, we derive 6 Thiessen polygon shapefiles and 6 Centroid shapefiles in this case. Click the symbols of 6 Thiessen polygon layers, choose no color of the Fill Color and choose the outline colors from blue to red. Check if the numbers of Thiessen polygons are correct: in this case, 5990+907+150+30+5+2=7084. The sum should equal to the number of NaturalCities_Centroid.



4. Creating the complex network

Before we create the complex network, we should check the field structures of the Thiessen and Centroid series. The Thiessen polygon shapefiles should at least have ‘Input_FID’ and ‘NodeCount’ columns. And the output Centroid shapefiles should have ‘NodeCount’ and ‘ORIG_FID’ columns. The ‘Input_FID’ and ‘ORIG_FID’ are the same meaning, indicating the original FID in the NaturalCities_Centroid shapefile. Create a new folder under your present working folder name it ‘graphdata’, copy 6 Thiessen and 6 Centroid shapefiles in that folder.



Table

Thiessen1

FID	Shape *	Shape_Leng	Shape_Area	Input_FID	NodeCount	ORIG_FID
0	Polygon ZM	1474.384457	138014.159771	3102	3	3102
1	Polygon ZM	2086.586182	291549.542494	3064	3	3064
2	Polygon ZM	2421.657169	369546.586019	6722	3	6722
3	Polygon ZM	2209.091401	298844.295026	6678	6	6678

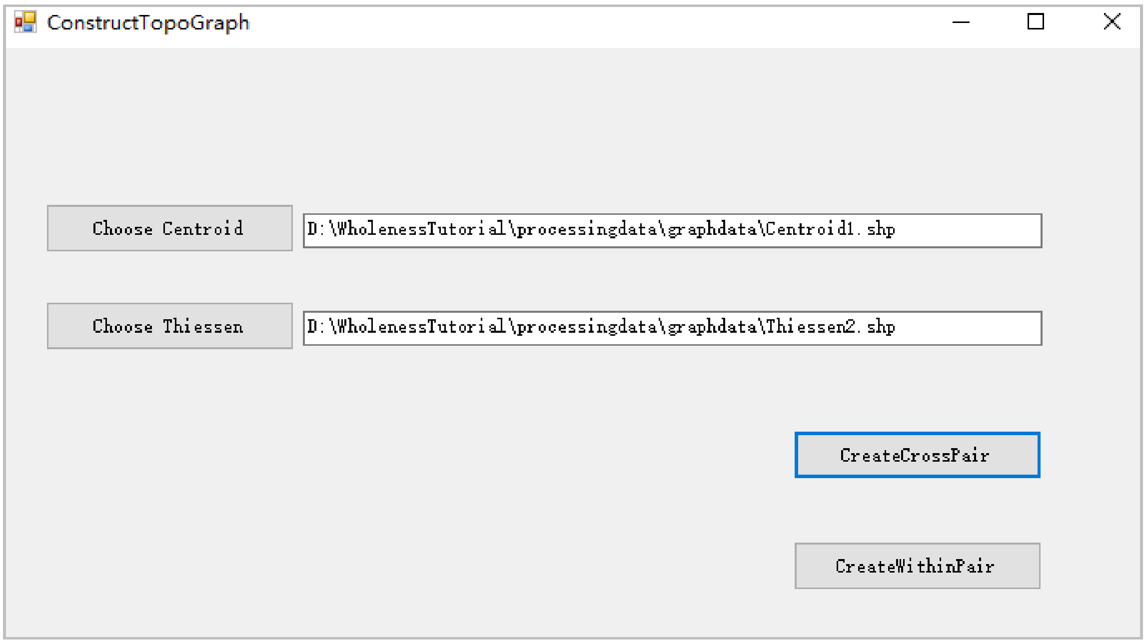

Table

Centroid1

FID	Shape *	NodeCount	ORIG_FID
0	Point ZM	8	0
1	Point ZM	7	1
2	Point ZM	13	2
3	Point ZM	8	3

1. Now we can take use of the TopoGraph.exe (see section 2.2 (3)) to create the graph. Open the TopoGraph.exe >> Click ‘Choose Centroid’ >> choose the Centroid1.shp in the ‘graphdata’ folder. >> Click ‘Choose Thiessen’ >> choose the next level Thiessen2.shp from the same folder >> Click ‘CreateCrossPair’ button >> Click OK when prompt windows pop up.

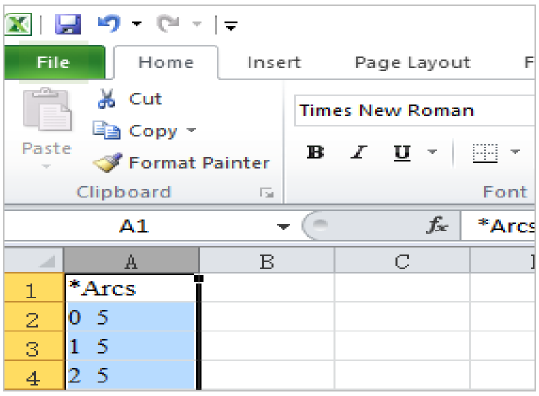
2. Repeat the previous step until Centroid5.shp and Thiessen6.shp have been processed in this case. You can find out the cross.txt file has been save in the same folder.

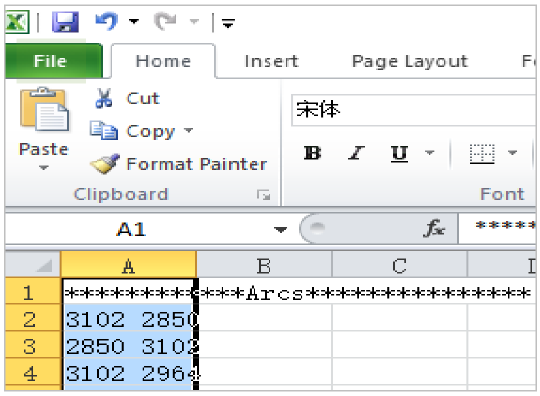


3. In order to create the within.txt, just follow the similar process, but only choose Thiessen1.shp to Thiessen6.shp each time and press the ‘CreateWithinPair’ button. Finally, the within.txt was automatically saved in your folder.

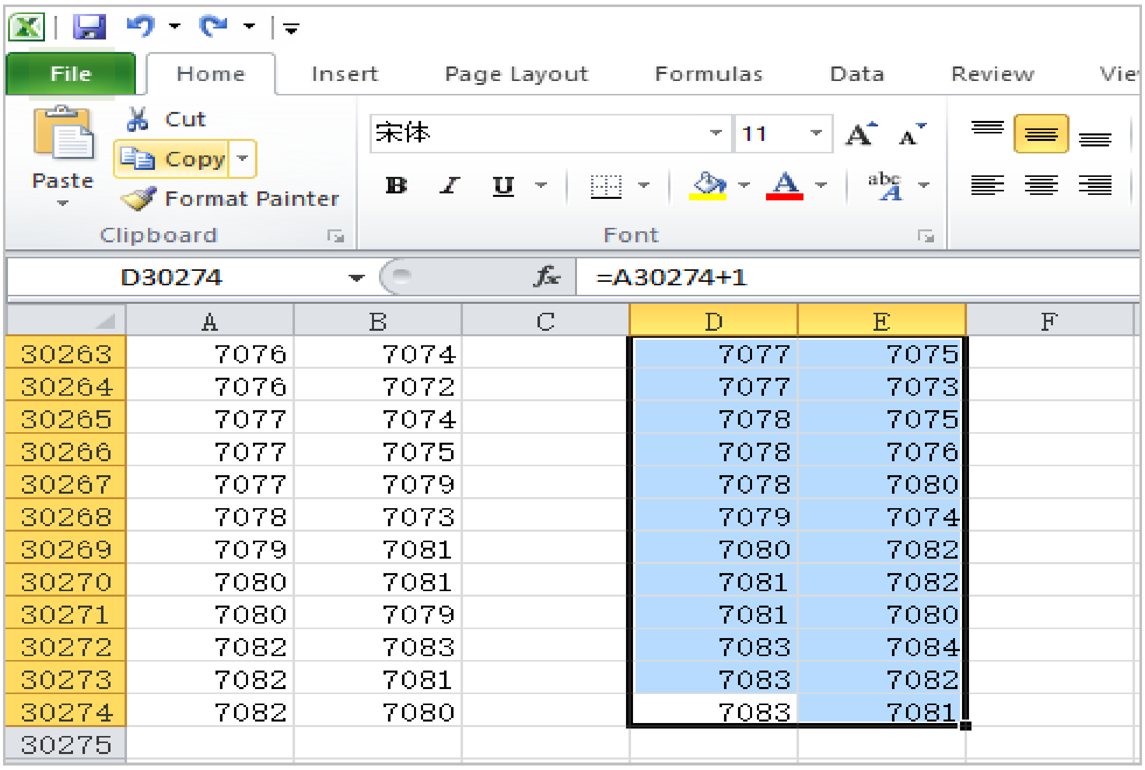
In the cross.txt file and within.txt file, there are respectively five and six groups of number pairs with ‘Arcs’ and headers, each number stands for the FID of the Centroid points in the NaturalCities_Centroid shapefile. Within each number pair, the first number stands for the start point of the network link and the second number is the end point. In this way, those number pairs consist of a large network indicating the topological relationship based on Thiessen polygons. In the following step, we sort these two text files in excel and generate an input text file for Matlab to calculate the degree of wholeness.

4. Open excel, copy the cross and within number pairs to two different sheets. In each sheet: Data >> Text to Colomns >> Delimited with space >> Sort smallest to largest >> Expand the selection area (Important!) >> Delete the text, such as ‘*Arcs’ and ‘***Arc***’ >> Copy and paste the number pairs of two sheets into one column.





In the Excel, plus one to each number, and copy the result pair to a text file, name it 'graphpair.txt'.



	A	B	C	D	E	F
30263	7076	7074		7077	7075	
30264	7076	7072		7077	7073	
30265	7077	7074		7078	7075	
30266	7077	7075		7078	7076	
30267	7077	7079		7078	7080	
30268	7078	7073		7079	7074	
30269	7079	7081		7080	7082	
30270	7080	7081		7081	7082	
30271	7080	7079		7081	7080	
30272	7082	7083		7083	7084	
30273	7082	7081		7083	7082	
30274	7082	7080		7083	7081	
30275						

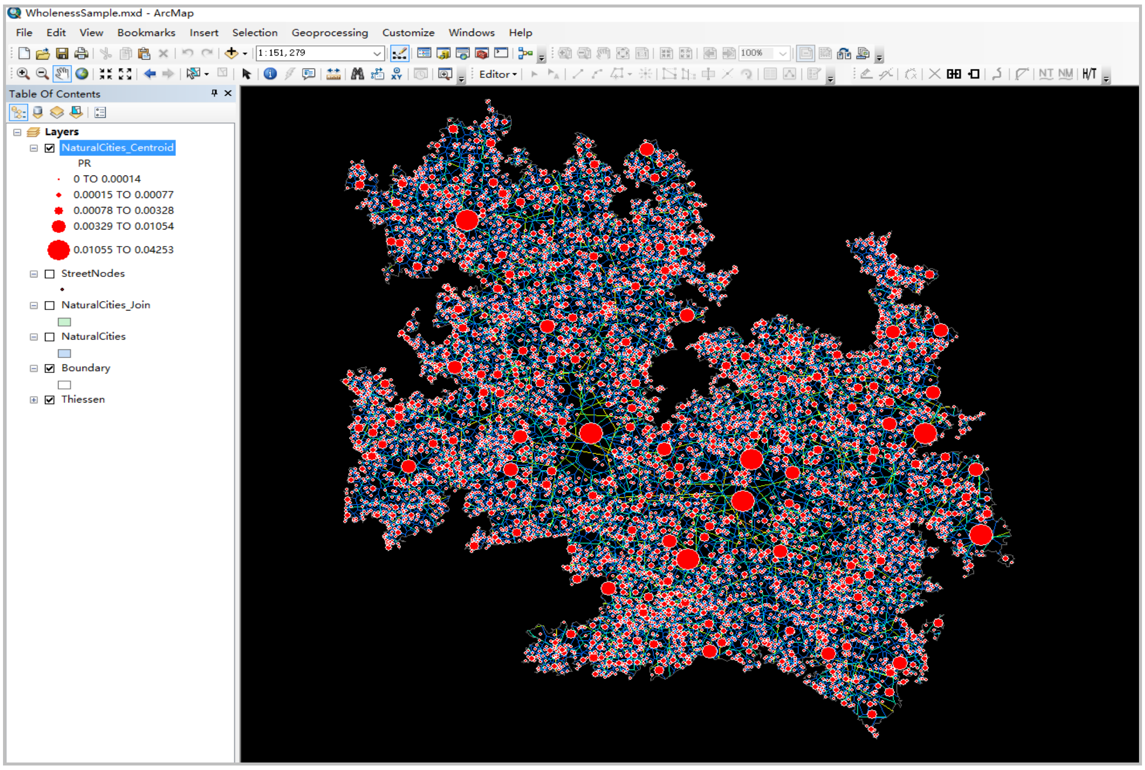
5. Calculating the degree of wholeness

Once we derive the graphpair.txt file, we can calculate the degree of wholeness in Matlab. The degree of wholeness of individual points can be measure by Pagerank (Jiang, 2015c). The Pagerank value indicates the importance of certain node inside a network structure. One node with high Pagerank is pointed by many other important points within a network. It can be applied not only in the web, but also in other areas such as biology, physics and road networks (Gleich, 2016). The Matlab code provided by David Gleich is used for calculating the Pagerank value. Note that you should check your input text file again to make sure that there are only number pairs in it. The input of the Pagerank code is a matrix with the link information of the network.

1. Open Matlab >> Put the input **graphpair.txt** and the **pagerank.m** file in the same folder >> Set current directory as the working folder >> In the command window, input the following code:

```
x=load('graphpair.txt');
G=sparse(7084,7084,0);
for e=1:size(x)
n1=x(e,1);
n2=x(e,2);
G(n1,n2)=1;
end
p=pagerank(G);
fid=fopen('graphPR.txt','wt');
for i=1:size(p,1)
fprintf(fid,'%d',p(i,:));
fprintf(fid,'\n');
end
```

2. The result text file ‘graphPR.txt’ is saved in the same folder. Add a new field called ‘PR’ to the NaturalCities_Centroid shapefile, the data type should be set as ‘double’ >> Add ‘graph.txt’ from the ArcCatalog to the map >> Right click to export the data as a dbf table >> name it as ‘PR.dbf’ >> Join this table to the NaturalCities_Centroid >> Calculate the PR based on ‘NaturalCities_Centroid.PR = PR.Field1’ >> Click HT mapping button of Axwoman >> Choose the value PR to create the final map. The result is shown as follows:

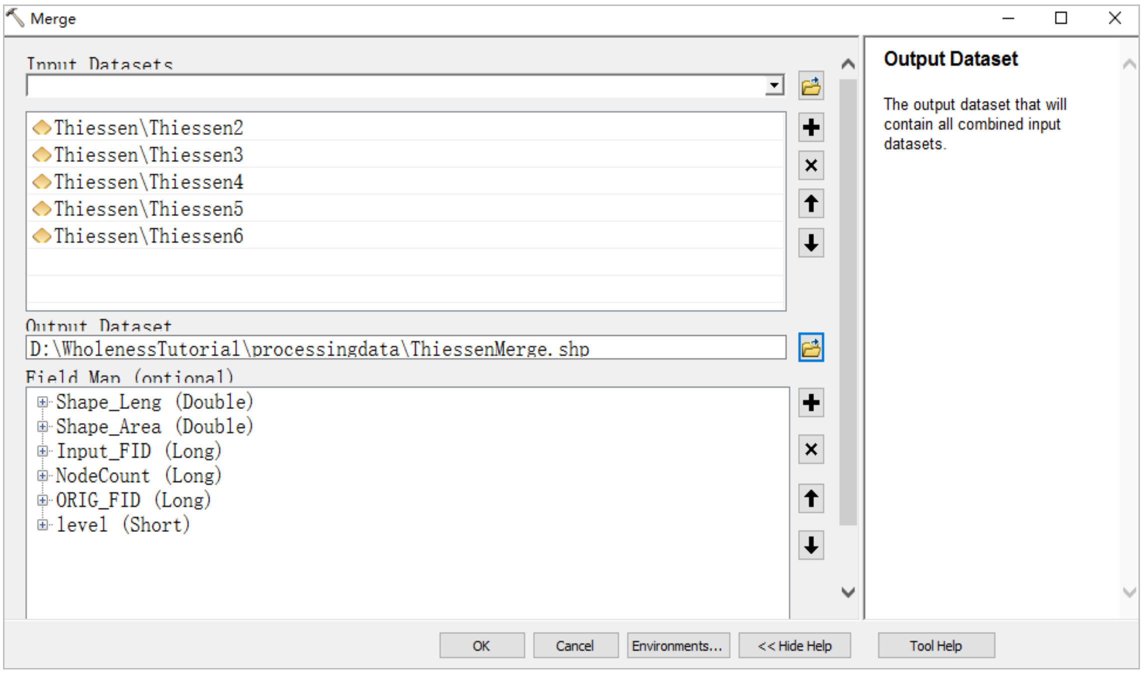


6. Radial layout visualization

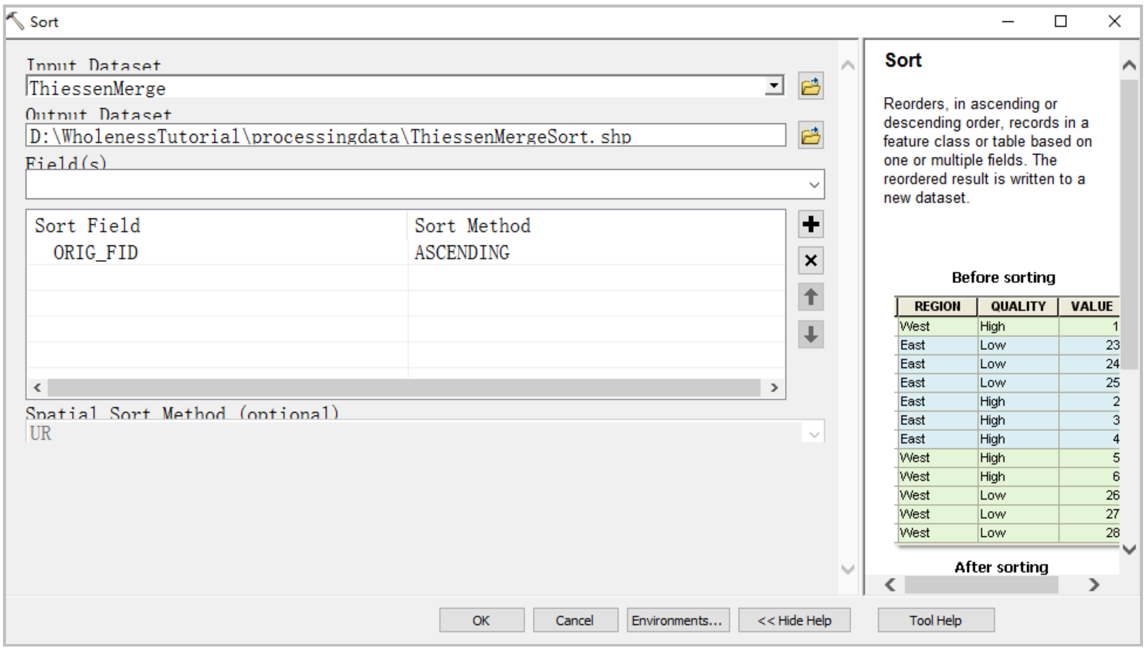
Once the centroids of the natural cities and their corresponding Thiessen polygons are derived, we can visualize the hierarchical level of the network in a radial layout using the topological representation radial layout program (see section 2.2 (5)). Before we start the program, we need to do some preparations on our previous data in ArcMap. There are 7084 centroids in our shapefile and it takes more than 20 minutes to generate the result. We select the first head part of those centroids using head/tail breaks for convenience.

1. Right click NaturalCities_Centroid layer >> Right click column ‘NodeCount’ >> Statistics >> Check the mean value is 13.716544 >> Select by attributes where [NodeCount > 13.716544] >> Right click the layer >> Export data >> name the shapefile ‘NaturalCities_CentroidTOP5.shp’.

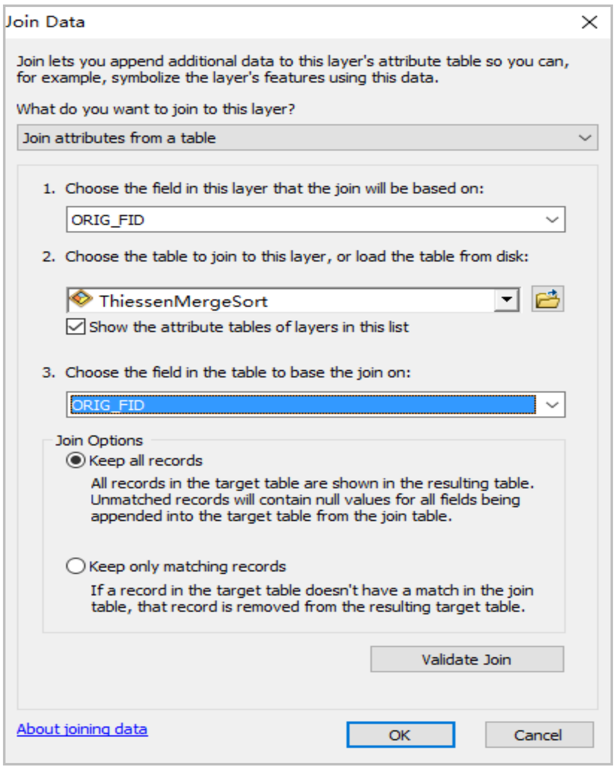
2. Add a new short integer field ‘level’ to the NaturalCities_CentroidTOP5.shp >> In the same way, add level fields to 5 Thiessen shapefiles (from Thiessen2.shp to Thiessen6.shp) >> In Thiessen2.shp, use Field Calculator to calculate the level field value as 4; Thiessen3.shp level as 3; Thiessen4.shp level as 2; Thiessen5.shp level as 1 and Thiessen6.shp level as 0 >> Geoprocessing >> Merge these five Thiessen shapefiles to one shapefile call ‘ThiessenMerge.shp’

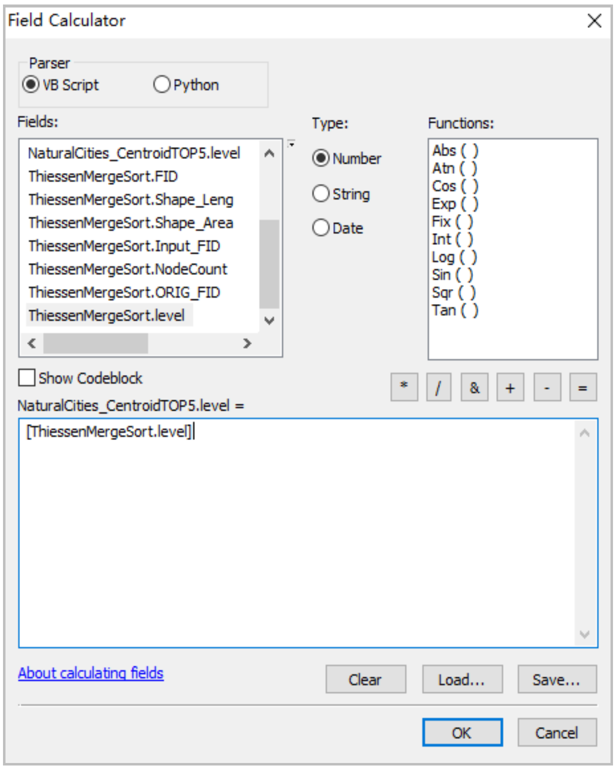


3. Open ArcToolbox >> Data management >> General >> Sort >> Choose the ThiessenMerge.shp as Input Datasets >> Sort Field is ORIG_FID or Input_FID >> Sort Method is Ascending >> Name the output as 'ThiessenMergeSort.shp' >> Click OK to start.

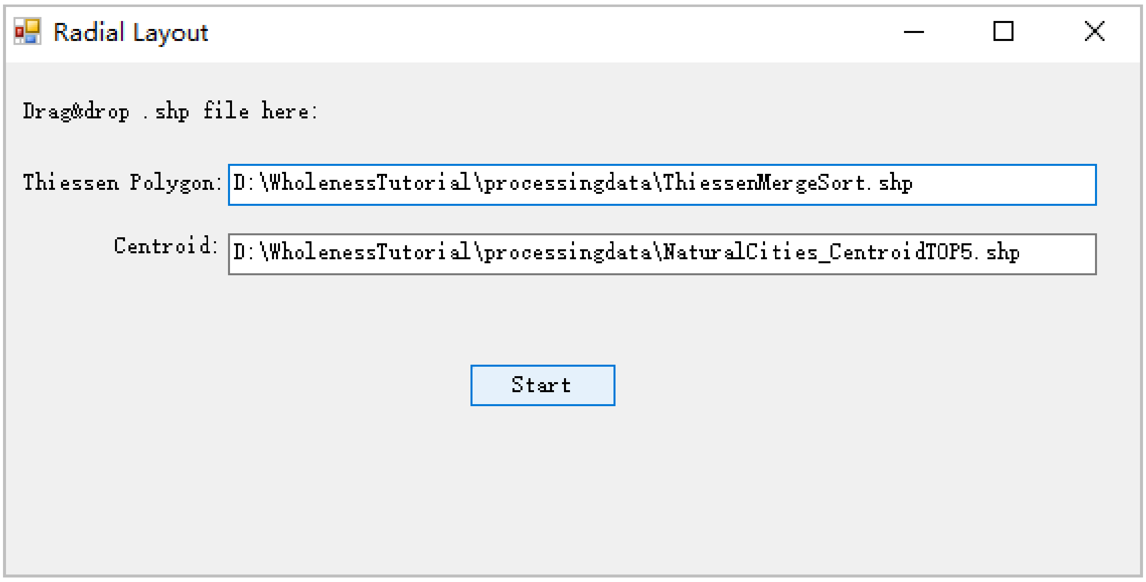


4. Right click NaturalCities_CentroidTOP5.shp >> Open Attribute Table >> Join and Relates >> Join the ThiessenMergeSort.shp to this layer >> Right click 'level' column >> Use Field Calculator to calculate the NaturalCities_CentroidTOP5.level = ThiessenMergeSort.level >> Remove all joins after calculation.



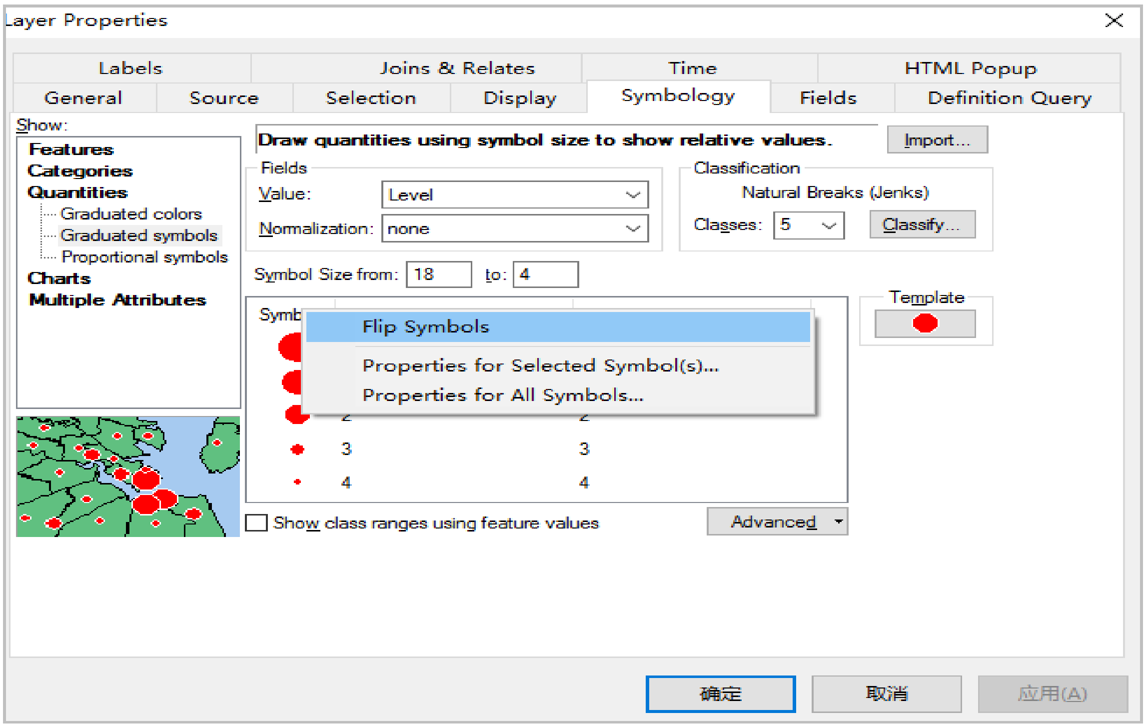


5. Open TopoRepr_10Dot2.exe or TopoRepr_10Dot0.exe >> Drag ThiessenMergeSort.shp to the Thiessen polygon textbox and drag NaturalCities_CentroidTOP5.shp to the Centroid textbox >> Click Start button.



6. The results including the node.shp and link.shp will be automatically saved in the same folder as the input shapefiles. Open ArcMap >> Insert a new data frame called 'Layout' >> Under this data frame add those two shapefiles >> Right click node layer >> Properties >> Symbology >> Quantities >> Choose Graduate symbols >> Flip symbols >> You can also customize your own symbols by clicking template >> Click the symbol of the link layer, you can customize the link layer.

7. Click the lower left corner of the viewer to switch to the Layout View, here you can arrange to data frames the way you prefer. Once you are satisfied with your final result, you can click File >> Export map to export your final map to a specific format you need.



6. Topological representation result

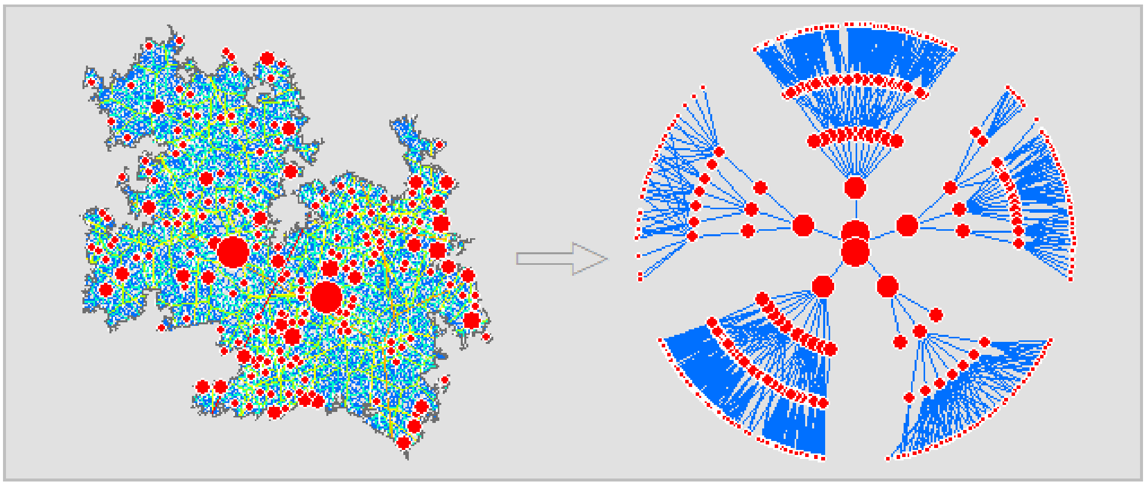


Figure 6.1 The topological representation of natural cities and radial layout visualization

References

Jiang B. (2013), Head/tail breaks: A new classification scheme for data with a heavy-tailed distribution, *The Professional Geographer*, 65 (3), 482 – 494.

Jiang B. (2015a), Wholeness as a hierarchical graph to capture the nature of space, *International Journal of Geographical Information Science*, 29(9), 1632–1648.

Jiang B. (2015b), Axwoman 6.3: An ArcGIS extension for urban morphological analysis, http://fromto.hig.se/~bjg/Axwoman/, University of G ävle, Sweden.

Jiang B. and Miao Y. (2015c), The evolution of natural cities from the perspective of location-based social media, *The Professional Geographer*, 67(2), 295 - 306.

Jiang B. (2016), A topological representation for taking cities as a coherent whole, https://www.researchgate.net/publication/305638074_A_Topological_Representation_for_Taking_Cities_as_a_Coherent_Whole

Jiang B. (2015c), Wholeness as a hierarchical graph to capture the nature of space, *International Journal of Geographical Information Science*, 29(9), 1632–1648.

Gleich D. F. (2016), Pagerank beyond the web, *SIAM Review*,57(3), 321-36.