Tutorial for Analyzing the scaling structure of natural streets

Chris de Rijke

Faculty of Engineering and Sustainable Development, Division of GIScience
University of Gävle, 801 76 Gävle, Sweden
E-mail: caderijke@gmail.com

In this tutorial a step by step guide is provided which shows the process of analyzing the scaling structure observed in natural streets. A set of natural streets will be provided, and its scaling structure will be analysed by performing head/tail breaks, visualizing the natural streets according to a head/tail break classification and finally analysing its connectivity by looking at its powerlaw distribution.

I. Introduction

The formation of natural streets is self-organized, naturally from individual road segments, based on the Gestalt principle of good continuity. Each segment joins with its neighbor according to some principles, such as every-best-fit, self-best-fit, and self-fit (Jiang, Zhao, and Yin 2008). The difference between these three principles is whether each segment considers every segment or only itself when choosing one (best) fit neighbor. For joining each segment with its most suitable neighbor segment, it is necessary to repeat the process of choosing the smallest deflection angle until the angle becomes greater than a given value. The result represents streets that are connected to each other and form a network topology. Natural streets differ from named streets that are identified by the unique names. The former (natural streets) has no problems with losses or errors of names, while named streets are difficult to implement due to the incompleteness of names. The concept of natural streets is widely used, such as computing the fewest-turn map directions or routes (Jiang and Liu 2011).

The previous study by Jiang (2009) has illustrated the fact that urban streets demonstrate the scaling property, which refers to far more less-connected than well-connected. This tutorial uses two methods to analyze scaling structure of natural streets. One method is power law distribution which is a type of heavy-tail distribution that can present the scaling pattern. The other is head/tail breaks (Jiang 2013) which is a classification scheme that divides data with heavy-tail distribution into the head (which has very few of large values) and the tail (which has a lot of small values). Data are classified around their arithmetic mean or average, recursively divided into head (above the mean) and tail (below the mean) by the use of head/tail breaks. This tutorial is divided into three more sections, section II is a data description and shows software requirements. Section III explains how to use head/tail breaks on natural streets and how you can visualize them. Finally section IV shows a powerlaw analysis on the examined set of natural streets.

II. Data resource and software support

This tutorial is provided with a set of natural streets from the city Rome which you can do yourself by following the full tutorial. (http://giscience.hig.se/binjiang/Axwoman/TutorialNaturalStreets14.pdf). The natural streets are generated with Axwoman software (http://giscience.hig.se/binjiang/axwoman/) and are based on freely available OpenStreetMap data (https://download.geofabrik.de/).

Necessary software for this tutorial:
Microsoft Excel (https://www.microsoft.com/en-us/microsoft-365/excel)
ArcMap (http://desktop.arcgis.com/en/arcmap/)
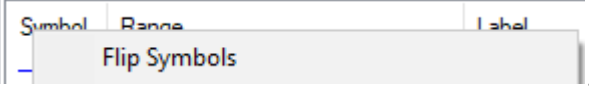Matlab (https://se.mathworks.com/downloads/web_downloads)
-    Matlab Powerlaw scripts (https://aaronclauset.github.io/powerlaws/)
Head/tail breaks calculator 1 column (https://github.com/ChrisdeRijke/HeadTailBreaksCalculator)

III. Scaling analysis using head/tail breaks

In this section the provided natural streets of Rome will be used to visualize the scaling structure inherently present in the streets.

1. First open the natural streets in ArcMap by clicking *Add data* and navigating to and selecting the *OSM_Rome_NS_3.shp* file. You should see a road network appear on your screen.
2. Next open the attribute table by *right-clicking* the layer in the Table of Contents and selecting *Open Attribute Table*.
3. Export the table to a .txt file so that we can read it in excel, by clicking the *Table options* button and selecting *Export*. Save the file as a text file so that it can be opened in excel later on. (NOT as a default file and personal geodatabase table).
4. Open the *ExcelHTcalculatorOneColumn50k.xlsx* head/tail breaks calculator file (see section II). Add a sheet and add the exported table from step 3 by going to the *Data* tab and selecting *From text* under *Get External Data*. The "*connect*" column is the one we are interested in as that indicates the amount of connections each natural street has with other streets. Select that whole column and input it into column A in the *results* sheet. A head/tail breaks classification will be calculated and shown. For this visualization you can choose to use head/tail breaks 1.0 or 2.0. The difference is that the most well connected streets (the absolute head) will show up clearer with head/tail breaks 2.0.
5. Take note of the HT index and the means at each row. The HT index will show how many classes you need in the classification while the means will show the break values at each class. In ArcMap *double-click* the natural street layer and go to *symbology*, under *quantities*, select *graduated colors*. Then select *Connect* as *Value* and select 7 classes for head/tail breaks 1.0 or 10 for head/tail breaks 2.0. Then click the *classify* button and input the correct *break values* as denoted in the excel calculator. (Note that you can copy paste the values one by one from the excel file).
6. Then select ▭ as color ramp and flip the colors by clicking on the symbol button and selecting *flip symbols*

Symbol	Range		Label
	Flip Symbols		

   . This is so that few connections are denoted by cold colors and many connections with warmer colors.

IV. Examining power law distribution

1. MATLAB software is used to examine if the degree distribution of the street network fits a power law distribution. For this the connectivity calculated in the previous section can be used. If you have already exported the Connectivity column for the head/tail breaks calculation you can use this again, otherwise follow step 2b from the previous section to export the column.
2. Start MATLAB and click *Browse for folder* to connect the folder you are working with.

Note: Make sure the degree vector file and Clauset MATLAB code files (Clauset, 2009) plfit.m, plplot.m and plpva.m are in this folder. (Go to https://aaronclauset.github.io/powerlaws/ to download these three MATLAB code files.)

a. Use Notepad to open the text file and delete the first row and save it as a text file. So that there are only connectivity numbers.
b. In the command window, type the code: *x=importdata('AvignonConnectivity.txt')* and press *Enter* on the keyboard to import the degree data of Avignon street network.

Command Window
```
>> x=importdata('AvignonConnectivity.txt')

x =

    2
    5
   11
   42
    1
    3
   17
    3
    3
    2
   17
    8
    2
    4
    2
    3
    2
    3
```

c. Type: *[alpha,xmin,D]=plfit(x)* and press *Enter*. This step is intended to get the alpha and minimum value using a power law fit function.

Command Window
```
    3
    2
    1
    2

>> [alpha,xmin,D]=plfit(x)

alpha =

    2.6700


xmin =

    3


D =

  -407.9459

>>
```

d. Next, use power law plot function, type *h=plplot(x,xmin,alpha)* and press *Enter*. A figure will display that shows a power law distributional model of Avignon street. The value along the x axis are degree values from the natural streets degree file (connectivity). The values along the y axis show the cumulative probability of obtaining those values. You can see specific X and Y values by hovering your mouse cursor over a blue dot.

Command Window
```
alpha =

    2.6700


xmin =

    3


D =

  -407.9459

>> h=plplot(x,xmin,alpha)

h =

    0.0001
    1.0001

>>
```
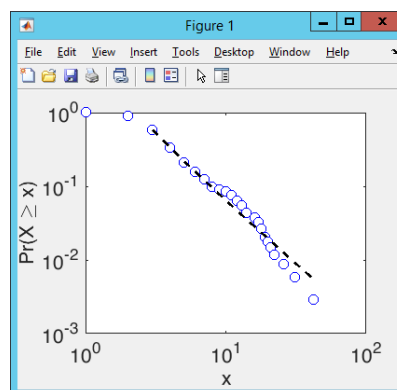
e. Calculate a p-value by using power law pva function. Type: *[p,gof]=plpva(x,xmin,'reps',500)* and press *Enter*.

Note: If the p-value is greater than 0.05, the data fits a power law distribution. For Avignon, the p-value is equal to 0.2180, which is greater than 0.05. (In contrast to traditional hypothesis testing in statistics, here we want high p values.) Note: Depending on how Avignon was prepared in pre-processing results may be different, however if the selected area is representative of Avignon it should represent a p-value greater than 0.05.

After completing above exercises, you have learned how to analyse natural streets, conduct head/tail breaks and examine power law distribution for data.

References:
Clauset A., Shalizi C. R., and Newman M. E. J. (2009), Power law distributions in empirical data, *SLAM Review*, 51, 661–703.
Jiang B., Zhao S., and Yin J. (2008), Self-organized natural roads for predicting traffic flow: a sensitivity study, *Journal of Statistical Mechanics: Theory and Experiment*, July, P07008.
Jiang B. and Liu X. (2011), Computing the fewest-turn map directions based on the connectivity of natural roads, *International Journal of Geographical Information Science*, 25(7), 1069–1082.
Jiang B. (2009), Street hierarchies: a minority of streets account for a majority of traffic flow, *International Journal of Geographical Information Science*, 23(8), 1033–1048.
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. (2019), *Axwoman in a Nutshell*,
https://www.researchgate.net/publication/337656401_Axwoman_in_a_Nutshell
Jiang B. (2019), A recursive definition of goodness of space for bridging the concepts of space and place for sustainability. *Sustainability*, 11(15), 4091.
Miao Y. (2012), *A Tutorial for Transforming a Line Shapefile to Pajek's Netfile, and then back to the shapefile*, University of Gävle, Sweden.
Sun X. (2012a), *Tutorial for Automatically Generating Axial Lines*, University of Gävle, Sweden.
Sun X. (2012b), *Small-world and Scale-free Properties of Urban Street Networks: A Tutorial*, University of Gävle, Sweden.
Wu S. (2014), *Tutorial for Applying Head/tail Breaks in ArcGIS*, University of Gävle, Sweden.