

Documentation

Plotting the nodes file with gnuplot

By

Muhammad Ali Hashmi

&

Markus Luczak-Roesch

CONTENTS

CONTENTS	ii
1 FileConverterGNUPlot	1
2 Gnuplot Plotting Script	1
2.1 Setting the terminal and fonts	1
2.2 Setting the Canvas Style	2
2.3 Set the Axis and Axis-Labels Formatting	2
2.3.1 Set x-axis data type	2
2.3.2 Set xtics and ytics	2
2.3.3 Set x- and y-axis labels	3
2.4 Set the Line Styles and Plot Formatting	3
2.5 Plot the Graph	4
3 Calling the gnuplot script	4

1 FileConverterGNUPlot

To plot the data with gnuplot (e.g. time-stamp vs entropy), the first step is to convert the “*nodes*” output file to simple text file containing two columns to be plotted on x and y-axis. That can easily be done with the scripts “**FileConverterGNUPlot.py**” and “**FileConverterGNUPlotTwitter.py**”. These two scripts are different because of the different timestamp formatting in twitter data. The scripts may be called as follows (taking the nodes-file in first argument):

```
➤ python3 FileConverterGNUPlot.py nodes-file.csv
```

2 Gnuplot Plotting Script

The script “stats_vs_time_png” is the gnuplot plotting script. The script is commented well to make it easily reproduceable. Here is the breakdown of the main plotting script.

2.1 Setting the terminal and fonts

Due to larger data points and very large eps output file, png terminal is chosen here. The size of the output png plot is 2420 X 4512 pixels which cannot be changed as desired. Different fonts can also be chosen. Transparent keyword makes the png file transparent. “*set output*” defines the output plot image file whose name can be changed here.

```
set terminal png size 2420,1512 enhanced font "Arial,28" transparent
set encoding utf8
set output 'output.png' # Define the output plot image.
```

2.2 Setting the Canvas Style

The canvas of the plot is defined below. “*set border*” sets a border around the plot with the border “*linewidth*” set as 2. This can be made thicker or thinner by changing its value. Apart from that, the margins can also be set for the plot using “*lmargin*” for left margin, “*rmargin*” for right side margin and so on.

```
set border linewidth 2
set lmargin at screen 0.10
set rmargin at screen 0.97
```

2.3 Set the Axis and Axis-Labels Formatting

The following line sets the formatting of the axes:

```
set format '%g'
```

The details of different formatting options can be found at http://gnuplot.sourceforge.net/docs_4.2/node184.html.

2.3.1 Set x-axis data type

The data type for x-axis can be specified whether the data is date/time, currency, etc. Also, its format can be set in the following piece of code:

```
set xdata time
set timefmt '"%Y-%m-%d %H:%M:%S"'
set format x '"%Y-%m-%d %H:%M:%S"'
```

Here in the first line the *xdata* is set as time. Then in the second line, time format is defined. In the third line, x-axis formatting is described.

2.3.2 Set xtics and ytics

In the code below, the x- and y-tics are defined.

```

### Set xtics below ###
set xtics font "Arial,20" offset -7,-5 scale 3 # Font for xtics
set xtics mirror in rotate by 45 autojustify

### Set ytics below ###
set ytics font "Arial,20" scale 3
set ytics mirror in

```

In the first lines, the font and font-size are defined. “*offset*” can be used to position the x-tics position at the axis. “*set xtics mirror in*” mirrors the x-axis and shows the small lines on the opposite side of the graph to make it easy to read from right side as well. The command “*rotate by 45*” rotates the x-tics by 45 degrees.

2.3.3 Set x- and y-axis labels

In the below pieces of codes for x- and y-axis, the first line selects the font and font-size for the axes, second line specifies the axis-label and its position, and the third line defines the range for the axis (i.e. minimum and maximum values).

```

### Set label for x-axis below ###
set xlabel font "Arial,40"
set xlabel 'Time Series' offset -1,-5
set xrange ["2003-01-01 00:00":"'2008-01-10 00:00"] # Set the range for X-axis

### Set label for y-axis below ###
set ylabel font "Arial,40"
set ylabel offset -1,1 'Entropy'
set yrange [1.0:14.70] # Set the range for Y-axis

```

2.4 Set the Line Styles and Plot Formatting

Default line styles available in gnuplot can be used but here are some personalized line styles which are used in this script. *lc* is short for “line colour”, *lt* is short for “line type”, *lw* is short for “linewidth”, *pt* for “point type”, and *ps* for “point size”.

```
set style line 1 linecolor rgb '#0060ad' linetype 1 linewidth 5 pointtype 13 pointsize 4 # blue
set style line 2 lc rgb '#dd181f' lt 1 lw 5 pt 2 ps 3 # red
set style line 3 lc rgb '#1b7e2d' lt 1 lw 5 pt 57 ps 3 # green
set style line 4 lc rgb 'black' lt 1 lw 5 pt 149 ps 3 # black
set style line 5 lc rgb '#8B008B' lt 1 lw 5 pt 28 ps 3 # magenta
```

2.5 Plot the Graph

Finally, the below line is responsible to plot the graph using all the settings defined above.

```
plot "output.txt" using 1:2 w lp ls 1 notitle
```

Here, “output.txt” is the file to plot, 1:2 means plot column 1 against 2 in the file, and *notitle* specifies to print no title.

3 Calling the gnuplot script

Calling the gnuplot script is very simple on command line or terminal.

```
gnuplot stats_vs_time_png.sh
```

It will produce the plot as “output.png” in the current directory.