UNIVERSITÉ DE LORRAINE | UFR MATHÉMATIQUES INFORMATIQUE MÉCANIQUE ET AUTOMATIQUE

M2 SSI: Threat Intelligence Project Report

DELECOUR Sébastien
RIGAUX Florent

Introduction

By looking at the paper from lab52.io concerning APT groups, which can be found here :   , we got interested in playing with these data available.

We decided to take the same dataset from ThaiCERT, written for MISP galaxy/cluster format, and available here: https://apt.thaicert.or.th/cgi-bin/aptgroups.cgi

This dataset is often updated.

Since Lab52.io did a lot of statistics about this dataset, we decided to cross data from ThaiCERT and some data from our production, concerning countries, taken from the World Bank dataset.

For each country listed in the ThaiCERT dataset, we inserted some economical and political information in a JSON file.

Here are the fields we decided to insert into our JSON for each country :

```
{
   "Name": "Afghanistan",
   "Alpha_3_Code": "AFG",
   "Regime" : "Islamic Republic",
   "Region" : "Asia",
   "GDP(B$)" : "19.2",
   "Growth of GDP between 2000 and 2010 (%)": "370",
   "Growth of GDP between 2010 and 2020 (%)": "1.08",
   "Military Expenditure (% of general government expenditure)":"226.9"
 },
```

Beware that the script was executed with root privilege ! Don’t forget to change the paths to your convenience.
We decided to put countries information files into a **.project/countries/** directory and information about APT actors into a **./project/actors/** directory. The script was placed into **./project/actors/** directory.

The project

During this project, we decided to provide some statistics that weren't available on the lab52 report. However, most of our production isn't significant enough to be considered as truth, since the most important and useful statistics were already computed by lab52.

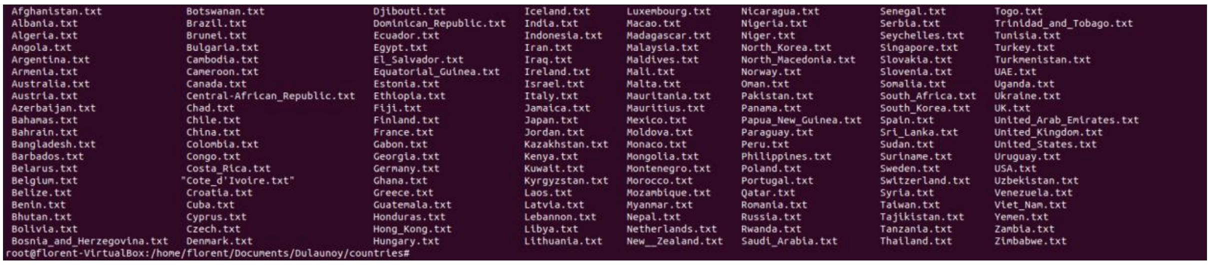
For our Country_list.json file, we used https://github.com/dominictarr/JSON.sh and its JSON.sh file to convert our json file into a txt one:

```
cat Country_list.json |./JSON.sh -l > Country_list.txt
```

Then we used the following command to create a folder for each of the 156 countries of our list:

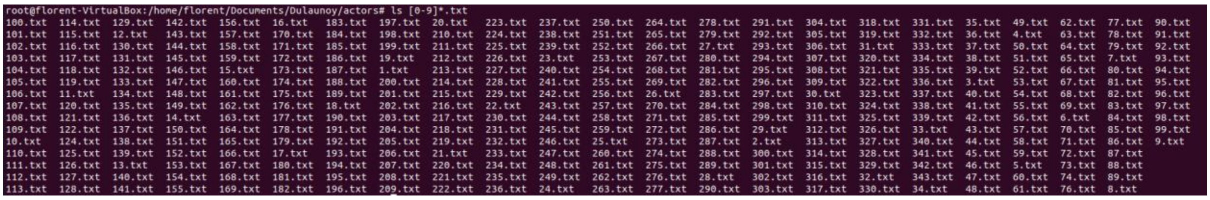
```
for i in $(seq 0 156);do grep "countries\",$i," Country_list.txt
> $(cat Country_list.txt | grep \"Name\" | sed -E
"s/^\[.*\].*\"([A-Z ?\'?a-zA-Z-]*)\"$/\1/" | sed -E "s/ /_/g" |
sed -n "$((i+1))  p").txt;done
```

We obtain:



```
 Afghanistan.txt              Botswanan.txt                  Djibouti.txt            Iceland.txt      Luxembourg.txt     Nicaragua.txt          Senegal.txt        Togo.txt
 Albania.txt                  Brazil.txt                     Dominican_Republic.txt  India.txt        Macao.txt          Nigeria.txt            Serbia.txt         Trinidad_and_Tobago.txt
 Algeria.txt                  Brunei.txt                     Ecuador.txt             Indonesia.txt    Madagascar.txt     Niger.txt              Seychelles.txt     Tunisia.txt
 Angola.txt                   Bulgaria.txt                   Egypt.txt               Iran.txt         Malaysia.txt       North_Korea.txt        Singapore.txt      Turkey.txt
 Argentina.txt                Cambodia.txt                   El_Salvador.txt         Iraq.txt         Maldives.txt       North_Macedonia.txt    Slovakia.txt       Turkmenistan.txt
 Armenia.txt                  Cameroon.txt                   Equatorial_Guinea.txt   Ireland.txt      Mali.txt           Norway.txt             Slovenia.txt       UAE.txt
 Australia.txt                Canada.txt                     Estonia.txt             Israel.txt       Malta.txt          Oman.txt               Somalia.txt        Uganda.txt
 Austria.txt                  Central-African_Republic.txt   Ethiopia.txt            Italy.txt        Mauritania.txt     Pakistan.txt           South_Africa.txt   Ukraine.txt
 Azerbaijan.txt               Chad.txt                       Fiji.txt                Jamaica.txt      Mauritius.txt      Panama.txt             South_Korea.txt    UK.txt
 Bahamas.txt                  Chile.txt                      Finland.txt             Japan.txt        Mexico.txt         Papua_New_Guinea.txt   Spain.txt          United_Arab_Emirates.txt
 Bahrain.txt                  China.txt                      France.txt              Jordan.txt       Moldova.txt        Paraguay.txt           Sri_Lanka.txt      United_Kingdom.txt
 Bangladesh.txt               Colombia.txt                   Gabon.txt               Kazakhstan.txt   Monaco.txt         Peru.txt               Sudan.txt          United_States.txt
 Barbados.txt                 Congo.txt                      Georgia.txt             Kenya.txt        Mongolia.txt       Philippines.txt        Suriname.txt       Uruguay.txt
 Belarus.txt                  Costa_Rica.txt                 Germany.txt             Kuwait.txt       Montenegro.txt     Poland.txt             Sweden.txt         USA.txt
 Belgium.txt                 "Cote_d'Ivoire.txt"             Ghana.txt               Kyrgyzstan.txt   Morocco.txt        Portugal.txt           Switzerland.txt    Uzbekistan.txt
 Belize.txt                   Croatia.txt                    Greece.txt              Laos.txt         Mozambique.txt     Qatar.txt              Syria.txt          Venezuela.txt
 Benin.txt                    Cuba.txt                       Guatemala.txt           Latvia.txt       Myanmar.txt        Romania.txt            Taiwan.txt         Viet_Nam.txt
 Bhutan.txt                   Cyprus.txt                     Honduras.txt            Lebannon.txt     Nepal.txt          Russia.txt             Tajikistan.txt     Yemen.txt
 Bolivia.txt                  Czech.txt                      Hong_Kong.txt           Libya.txt        Netherlands.txt    Rwanda.txt             Tanzania.txt       Zambia.txt
 Bosnia_and_Herzegovina.txt   Denmark.txt                    Hungary.txt             Lithuania.txt    New__Zealand.txt   Saudi_Arabia.txt       Thailand.txt       Zimbabwe.txt
root@florent-VirtualBox:/home/florent/Documents/Dulaunoy/countries#
```

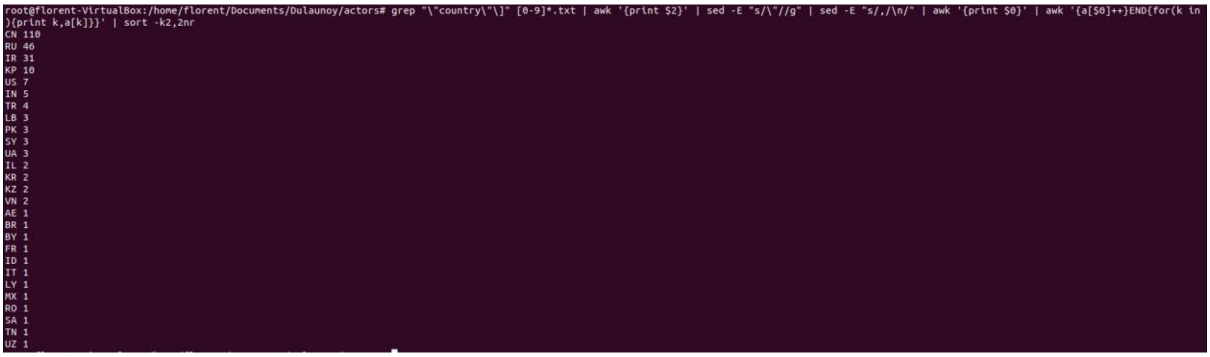
Then, we followed the tutorial provided by lab52 to extract each APT group and create a specific txt file for each of them. Those files were placed in **./project/actors/** directory, and have the particularity to be number entitled:



```
root@florent-VirtualBox:/home/florent/Documents/Dulaunoy/actors# ls [0-9]*.txt
100.txt  114.txt  129.txt  142.txt  156.txt  16.txt   183.txt  197.txt  20.txt   223.txt  237.txt  250.txt  264.txt  278.txt  291.txt  304.txt  318.txt  331.txt  35.txt  49.txt  62.txt  77.txt  90.txt
101.txt  115.txt  12.txt   143.txt  157.txt  170.txt  184.txt  198.txt  210.txt  224.txt  238.txt  251.txt  265.txt  279.txt  292.txt  305.txt  319.txt  332.txt  36.txt  4.txt   63.txt  78.txt  91.txt
102.txt  116.txt  130.txt  144.txt  158.txt  171.txt  185.txt  199.txt  211.txt  225.txt  239.txt  252.txt  266.txt  27.txt   293.txt  306.txt  31.txt   333.txt  37.txt  50.txt  64.txt  79.txt  92.txt
103.txt  117.txt  131.txt  145.txt  159.txt  172.txt  186.txt  19.txt   212.txt  226.txt  23.txt   253.txt  267.txt  280.txt  294.txt  307.txt  320.txt  334.txt  38.txt  51.txt  65.txt  7.txt   93.txt
104.txt  118.txt  132.txt  146.txt  15.txt   173.txt  187.txt  1.txt    213.txt  227.txt  240.txt  254.txt  268.txt  281.txt  295.txt  308.txt  321.txt  335.txt  39.txt  52.txt  66.txt  80.txt  94.txt
105.txt  119.txt  133.txt  147.txt  160.txt  174.txt  188.txt  200.txt  214.txt  228.txt  241.txt  255.txt  269.txt  282.txt  296.txt  309.txt  322.txt  336.txt  3.txt   53.txt  67.txt  81.txt  95.txt
106.txt  11.txt   134.txt  148.txt  161.txt  175.txt  189.txt  201.txt  215.txt  229.txt  242.txt  256.txt  26.txt   283.txt  297.txt  30.txt   323.txt  337.txt  40.txt  54.txt  68.txt  82.txt  96.txt
107.txt  120.txt  135.txt  149.txt  162.txt  176.txt  18.txt   202.txt  216.txt  22.txt   243.txt  257.txt  270.txt  284.txt  298.txt  310.txt  324.txt  338.txt  41.txt  55.txt  69.txt  83.txt  97.txt
108.txt  121.txt  136.txt  14.txt   163.txt  177.txt  190.txt  203.txt  217.txt  230.txt  244.txt  258.txt  271.txt  285.txt  299.txt  311.txt  325.txt  339.txt  42.txt  56.txt  6.txt   84.txt  98.txt
109.txt  122.txt  137.txt  150.txt  164.txt  178.txt  191.txt  204.txt  218.txt  231.txt  245.txt  259.txt  272.txt  286.txt  29.txt   312.txt  326.txt  33.txt   43.txt  57.txt  70.txt  85.txt  99.txt
10.txt   124.txt  138.txt  151.txt  165.txt  179.txt  192.txt  205.txt  219.txt  232.txt  246.txt  25.txt   273.txt  287.txt  2.txt    313.txt  327.txt  340.txt  44.txt  58.txt  71.txt  86.txt  9.txt
110.txt  125.txt  139.txt  152.txt  166.txt  17.txt   193.txt  206.txt  21.txt   233.txt  247.txt  260.txt  274.txt  288.txt  300.txt  314.txt  328.txt  341.txt  45.txt  59.txt  72.txt  87.txt
111.txt  126.txt  13.txt   153.txt  167.txt  180.txt  194.txt  207.txt  220.txt  234.txt  248.txt  261.txt  275.txt  289.txt  301.txt  315.txt  329.txt  342.txt  46.txt  5.txt   73.txt  88.txt
112.txt  127.txt  140.txt  154.txt  168.txt  181.txt  195.txt  208.txt  221.txt  235.txt  249.txt  262.txt  276.txt  28.txt   302.txt  316.txt  32.txt   343.txt  47.txt  60.txt  74.txt  89.txt
113.txt  128.txt  141.txt  155.txt  169.txt  182.txt  196.txt  209.txt  222.txt  236.txt  24.txt   263.txt  277.txt  290.txt  303.txt  317.txt  330.txt  34.txt   48.txt  61.txt  76.txt  8.txt
```

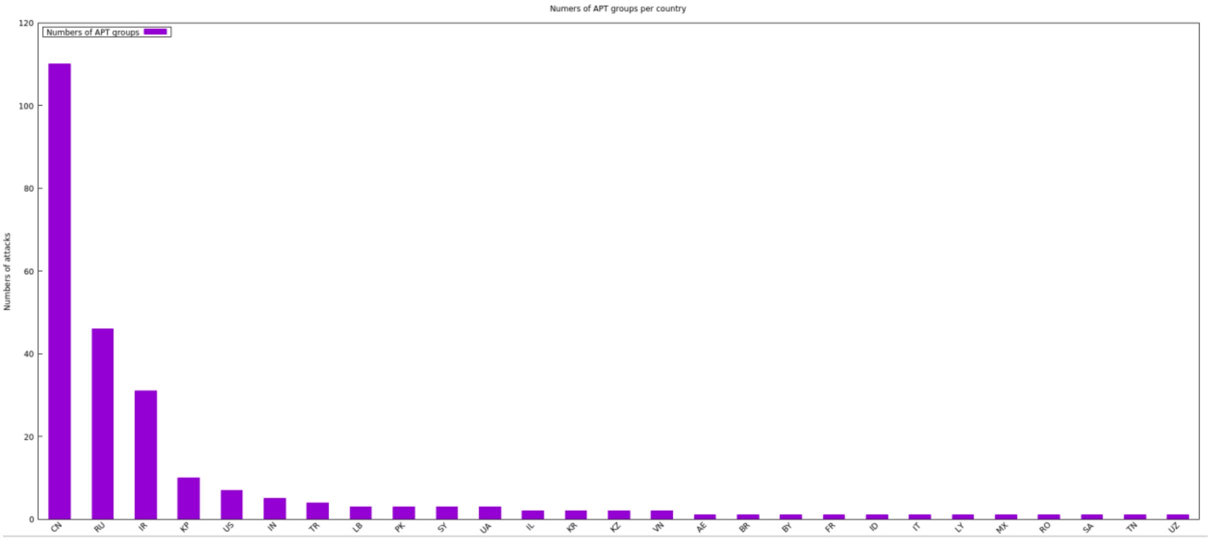
We removed files 123.txt and 75.txt because they only showed when UK GHCQ and US CIA were created, a long time ago.

Then we display the number of groups per native countries:



```
root@florent-VirtualBox:/home/florent/Documents/Dulaunoy/actors# grep "\"country\"\]" [0-9]*.txt | awk '{print $2}' | sed -E "s/\"//g" | sed -E "s/,/\n/" | awk '{print $0}' | awk '{a[$0]++}END{for(k in
){print k,a[k]}}' | sort -k2,2nr
CN 110
RU 46
IR 31
KP 10
US 7
IN 5
TR 4
LB 3
PK 3
SY 3
UA 3
IL 2
KR 2
KZ 2
VN 2
AE 1
BR 1
BY 1
FR 1
ID 1
IT 1
LY 1
MX 1
RO 1
SA 1
TN 1
UZ 1
```

We can export it in a file to display it with gnuplot:



We can then see that China, Russia and Iran are the champions, by having a lot of APT groups.

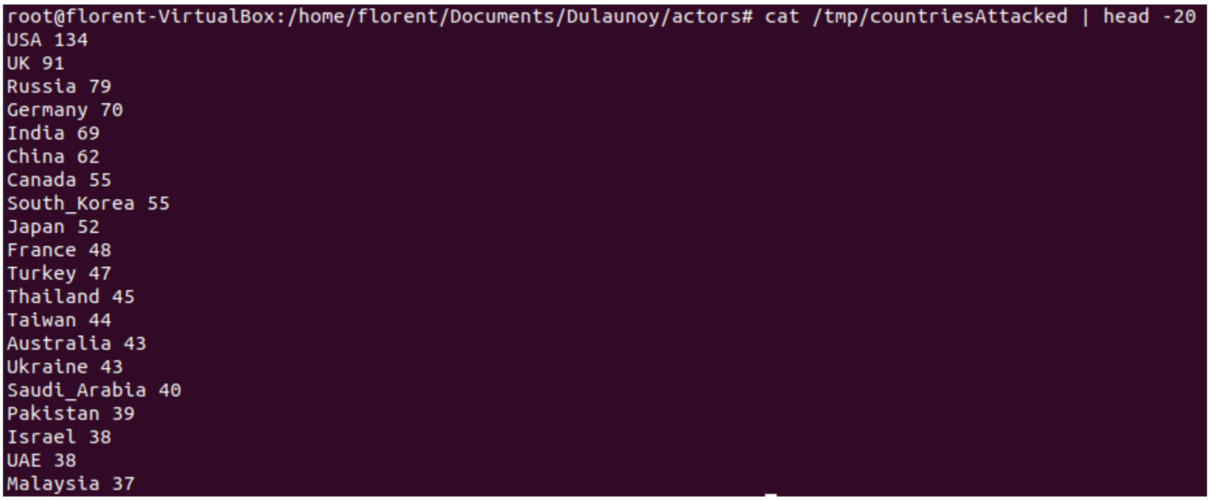
Let’s see which countries are the most attacked with commands:

```
grep -e “\”cfr-suspected-victims\”,[0-9]*” [0-9]*.txt | awk -F
‘”’ ‘{print $8}’ | grep -v “Unknown” | awk ‘{a[$0]++}END{for(k in
a){print k,a[k]}}’ | sort -k2,2nr > /tmp/countriesAttacked
```

We write the result into a file, which we will filter with:

```
Cat /tmp/countriesAttacked | sed -E “s/([A-Za-z]) ([A-Za-
z])/\1_\2/g” | sort -k2,2nr | sponge /tmp/countriesAttacked
```

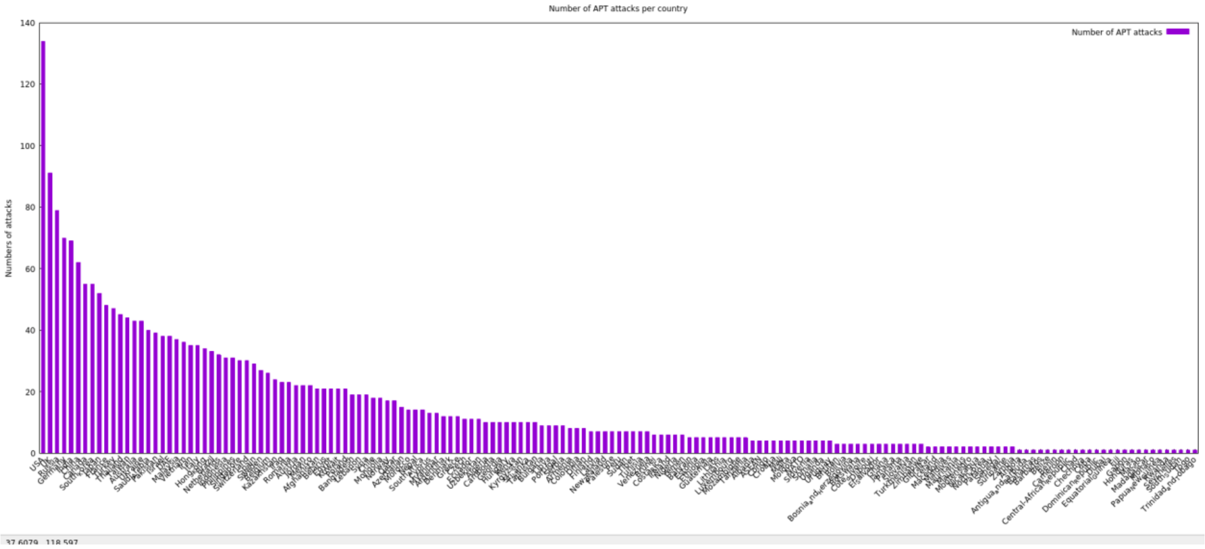
This allows to sort the file correctly and when we cat it, we obtain :



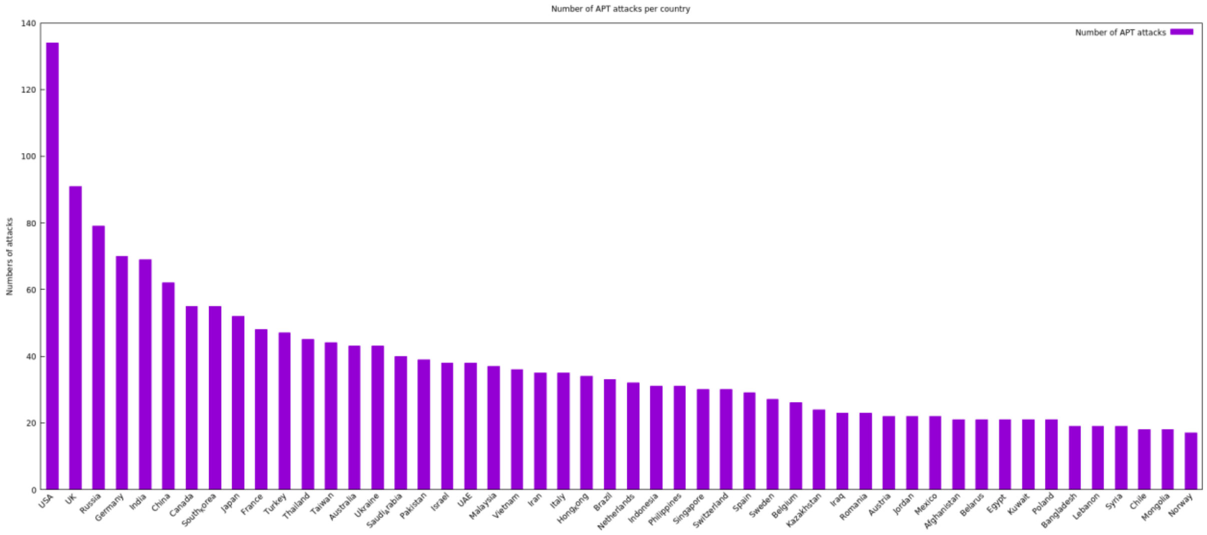
```
root@florent-VirtualBox:/home/florent/Documents/Dulaunoy/actors# cat /tmp/countriesAttacked | head -20
USA 134
UK 91
Russia 79
Germany 70
India 69
China 62
Canada 55
South_Korea 55
Japan 52
France 48
Turkey 47
Thailand 45
Taiwan 44
Australia 43
Ukraine 43
Saudi_Arabia 40
Pakistan 39
Israel 38
UAE 38
Malaysia 37
```

Let’s plot it to get more comfortable. We’ll use gnuplot and the following commands:

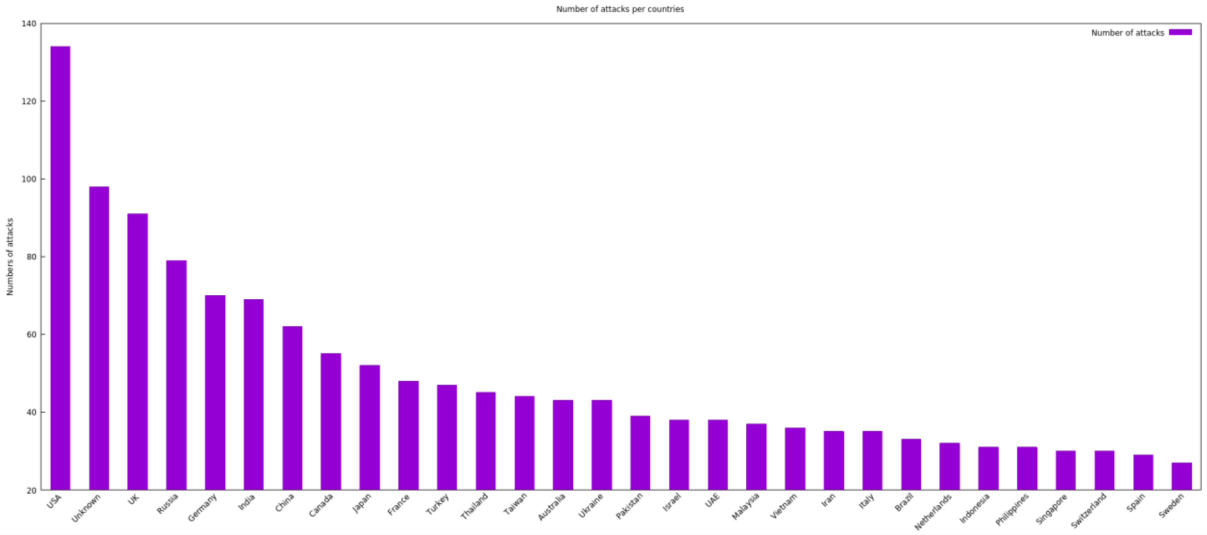
```
set boxwidth 0.5
set xtics rotate by 45 right
set style fill solid
set title “Number of APT attacks per country”
set tics nomirror
set ylabel “Numbers of attacks”
plot “/tmp/countriesAttacked” using 2:xticlabel(1) with boxes
title “Number of APT attacks”
```



It’s not really readable, so we’ll plot this for top 50 countries attacked:



It’s way better. We can now see that US, UK and Russia are the most targeted countries.



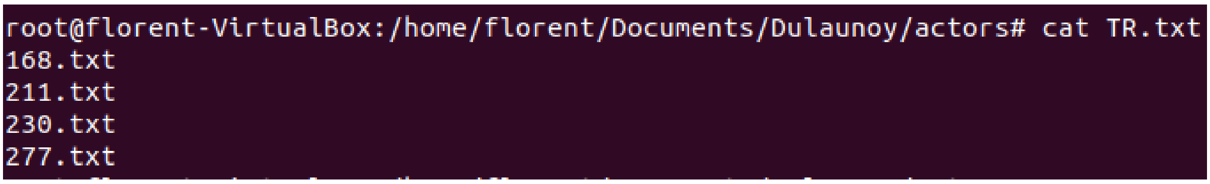
Another interesting thing is that we don’t know the targets for a lot of attacks, since “Unknown” arrives in second position.

Next thing we do is taking only countries which have at least 4 APT groups and sending the list into a file with the command:

```
grep "\"country\"\]" [0-9]*.txt | awk '{print $2}' | sed -E
"s/\"//g" | sed -E "s/,/\n/" | awk '{print $0}' | awk
'{a[$0]++}END{for(k in a){print k,a[k]}}' | sort -k2,2nr | awk
'{if ($2 >= 4) print $1}' > sampleA.txt
```

Then, we use the script **script.sh.** It allows us to parse every *[0-9]*.txt* file to put them into their respective country file.
For example, file 17.txt will be place into file CN.txt

Here are all the files related to Turkey:



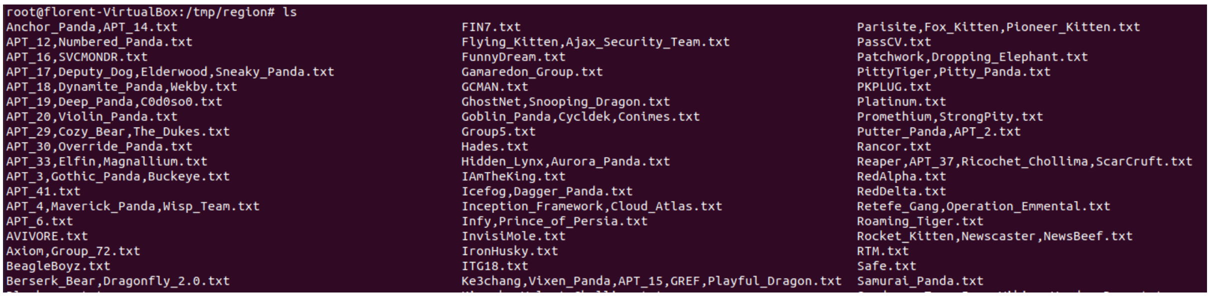
```
root@florent-VirtualBox:/home/florent/Documents/Dulaunoy/actors# cat TR.txt
168.txt
211.txt
230.txt
277.txt
```

Then, for every country file, we will read every group file, search the APT name, create a txt file with it. Afterwards, we copy the countries targeted by the APT group into the dedicated APT file, and also into a general /tmp/countries files. (We can note that some APT groups don’t have victims list. Then we added “Unknown” into its file, as their only victim)

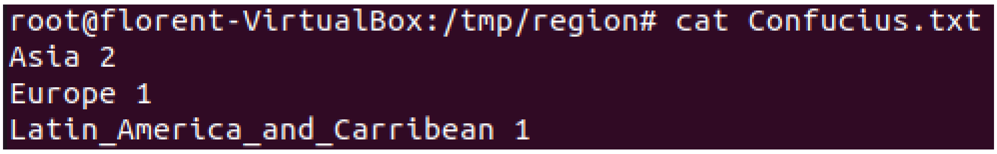
Once it’s done, we loop to read those countries. We then search for the specified field, like GDP, military budget, or budget; and we create a dedicated directory in the **/tmp/** directory, where we store information into a text file named with the APT group.

Once the script was executed, we have some dedicated files and repertories within /tmp/

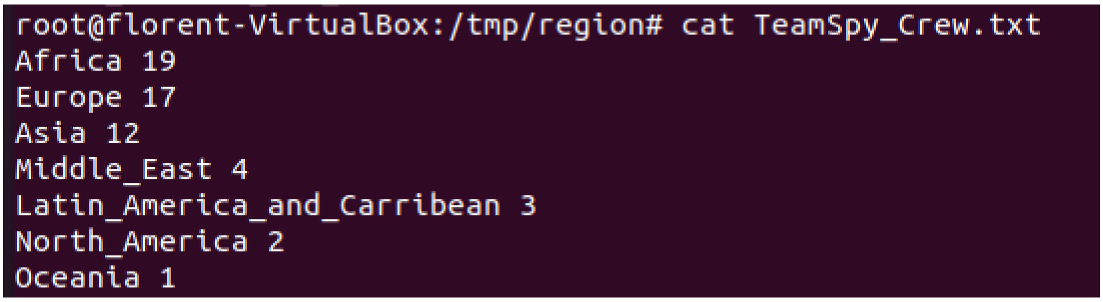
If we go to **/tmp/region/** we have the list of all APT groups and their targeted regions, referred by our Country_list file:



```
root@florent-VirtualBox:/tmp/region# ls
Anchor_Panda,APT_14.txt                                      FIN7.txt                                             Parisite,Fox_Kitten,Pioneer_Kitten.txt
APT_12,Numbered_Panda.txt                                    Flying_Kitten,Ajax_Security_Team.txt                 PassCV.txt
APT_16,SVCMONDR.txt                                          FunnyDream.txt                                       Patchwork,Dropping_Elephant.txt
APT_17,Deputy_Dog,Elderwood,Sneaky_Panda.txt                 Gamaredon_Group.txt                                  PittyTiger,Pitty_Panda.txt
APT_18,Dynamite_Panda,Wekby.txt                              GCMAN.txt                                            PKPLUG.txt
APT_19,Deep_Panda,C0d0so0.txt                                GhostNet,Snooping_Dragon.txt                         Platinum.txt
APT_20,Violin_Panda.txt                                      Goblin_Panda,Cycldek,Conimes.txt                     Promethium,StrongPity.txt
APT_29,Cozy_Bear,The_Dukes.txt                               Group5.txt                                           Putter_Panda,APT_2.txt
APT_30,Override_Panda.txt                                    Hades.txt                                            Rancor.txt
APT_33,Elfin,Magnallium.txt                                  Hidden_Lynx,Aurora_Panda.txt                         Reaper,APT_37,Ricochet_Chollima,ScarCruft.txt
APT_3,Gothic_Panda,Buckeye.txt                               IAmTheKing.txt                                       RedAlpha.txt
APT_41.txt                                                   Icefog,Dagger_Panda.txt                              RedDelta.txt
APT_4,Maverick_Panda,Wisp_Team.txt                           Inception_Framework,Cloud_Atlas.txt                  Retefe_Gang,Operation_Emmental.txt
APT_6.txt                                                    Infy,Prince_of_Persia.txt                            Roaming_Tiger.txt
AVIVORE.txt                                                  InvisiMole.txt                                       Rocket_Kitten,Newscaster,NewsBeef.txt
Axiom,Group_72.txt                                           IronHusky.txt                                        RTM.txt
BeagleBoyz.txt                                               ITG18.txt                                            Safe.txt
Berserk_Bear,Dragonfly_2.0.txt                               Ke3chang,Vixen_Panda,APT_15,GREF,Playful_Dragon.txt  Samurai_Panda.txt
```

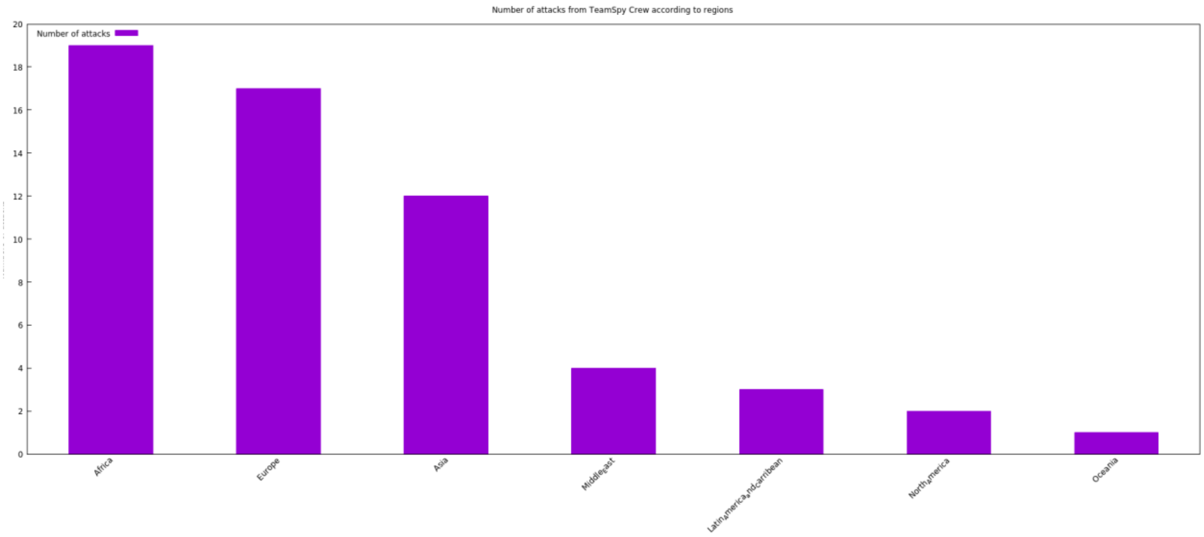


```
root@florent-VirtualBox:/tmp/region# cat Confucius.txt
Asia 2
Europe 1
Latin_America_and_Carribean 1
```

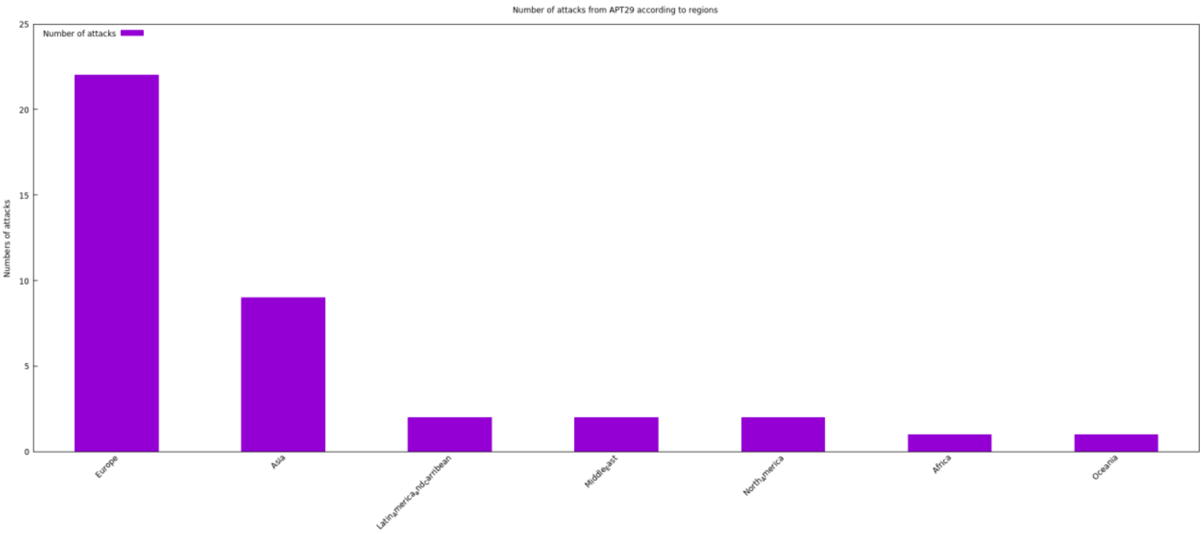


```
root@florent-VirtualBox:/tmp/region# cat TeamSpy_Crew.txt
Africa 19
Europe 17
Asia 12
Middle_East 4
Latin_America_and_Carribean 3
North_America 2
Oceania 1
```

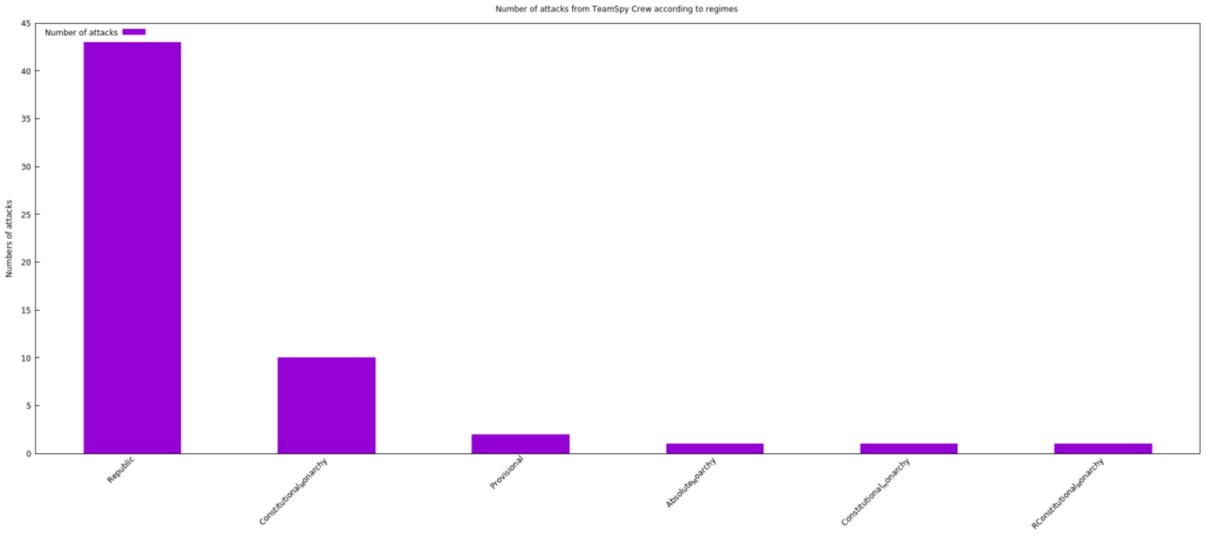
Let’s plot this last one:



It seems that it doesn’t target any specific region compared to APT29, which seems to like Europe:

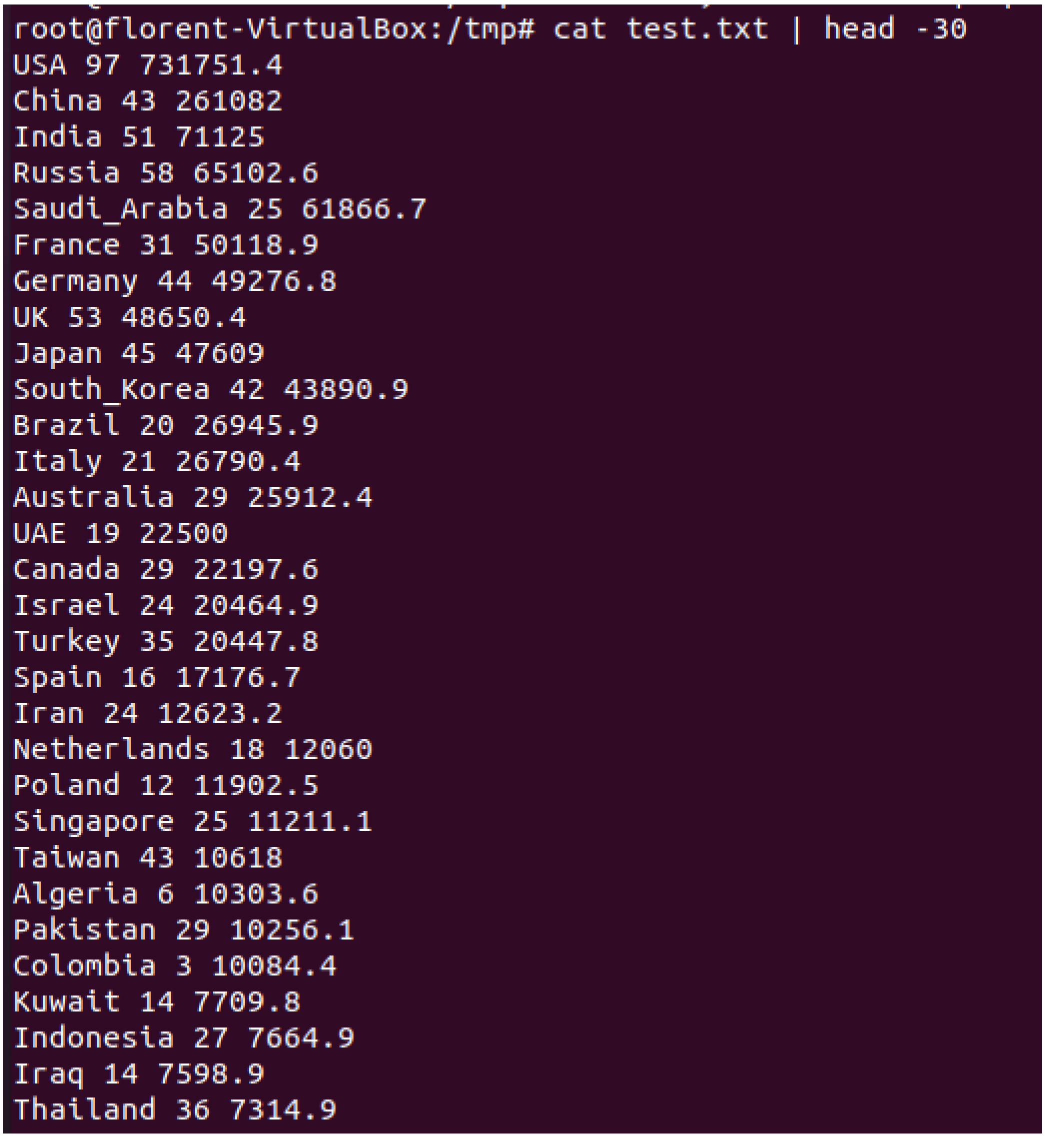


Concerning the regimes, TeamSpy Crew is a Russian group specialized in Information thief and espionage. Looking from a regime point of view:



We notice that it attacked mainly republics. Of course, it’s just statistics. They don’t particularly target republics.

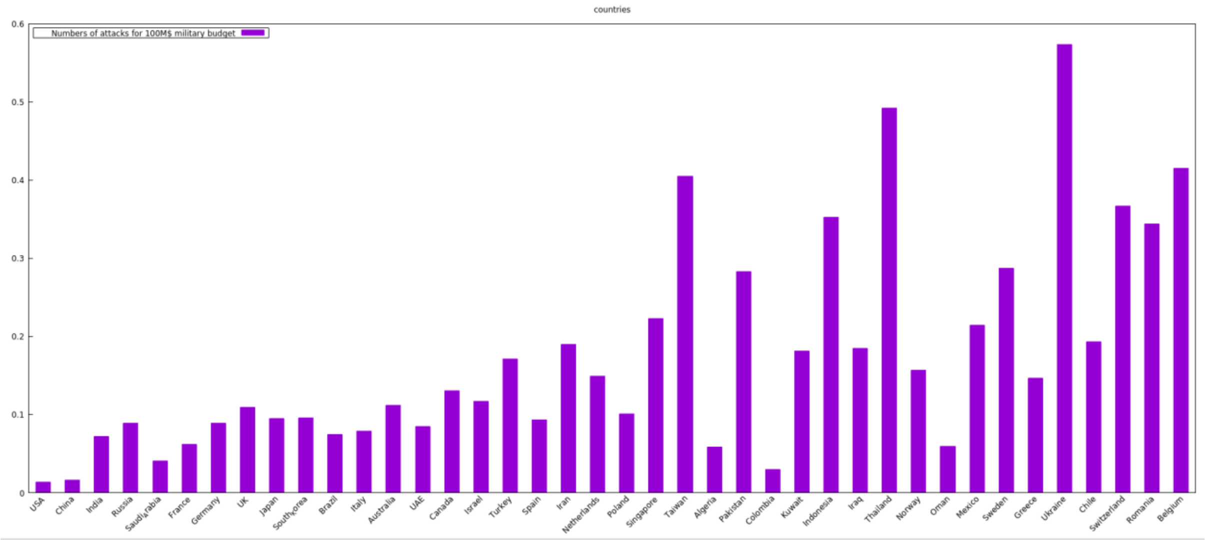
For **/tmp/military** file, if we display top 30 countries with biggest military budget, we have:



```
root@florent-VirtualBox:/tmp# cat test.txt | head -30
USA 97 731751.4
China 43 261082
India 51 71125
Russia 58 65102.6
Saudi_Arabia 25 61866.7
France 31 50118.9
Germany 44 49276.8
UK 53 48650.4
Japan 45 47609
South_Korea 42 43890.9
Brazil 20 26945.9
Italy 21 26790.4
Australia 29 25912.4
UAE 19 22500
Canada 29 22197.6
Israel 24 20464.9
Turkey 35 20447.8
Spain 16 17176.7
Iran 24 12623.2
Netherlands 18 12060
Poland 12 11902.5
Singapore 25 11211.1
Taiwan 43 10618
Algeria 6 10303.6
Pakistan 29 10256.1
Colombia 3 10084.4
Kuwait 14 7709.8
Indonesia 27 7664.9
Iraq 14 7598.9
Thailand 36 7314.9
```

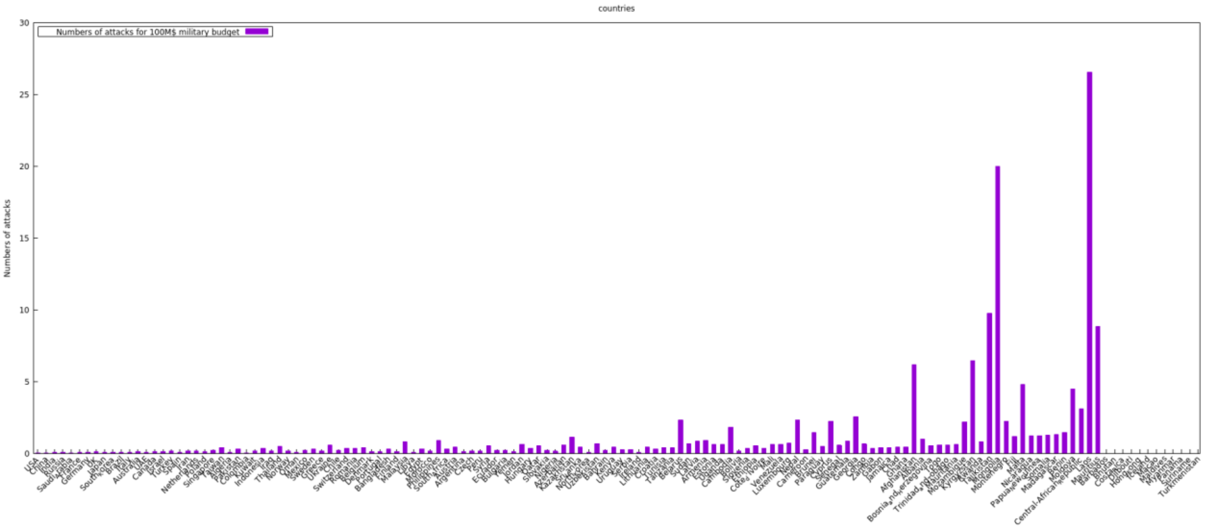
Second column corresponds to the number of attacks noticed. Third one corresponds to the military budget for the country.

Let’s plot the information concerning the number of attacks towards a country, for 100M$ to its military budget



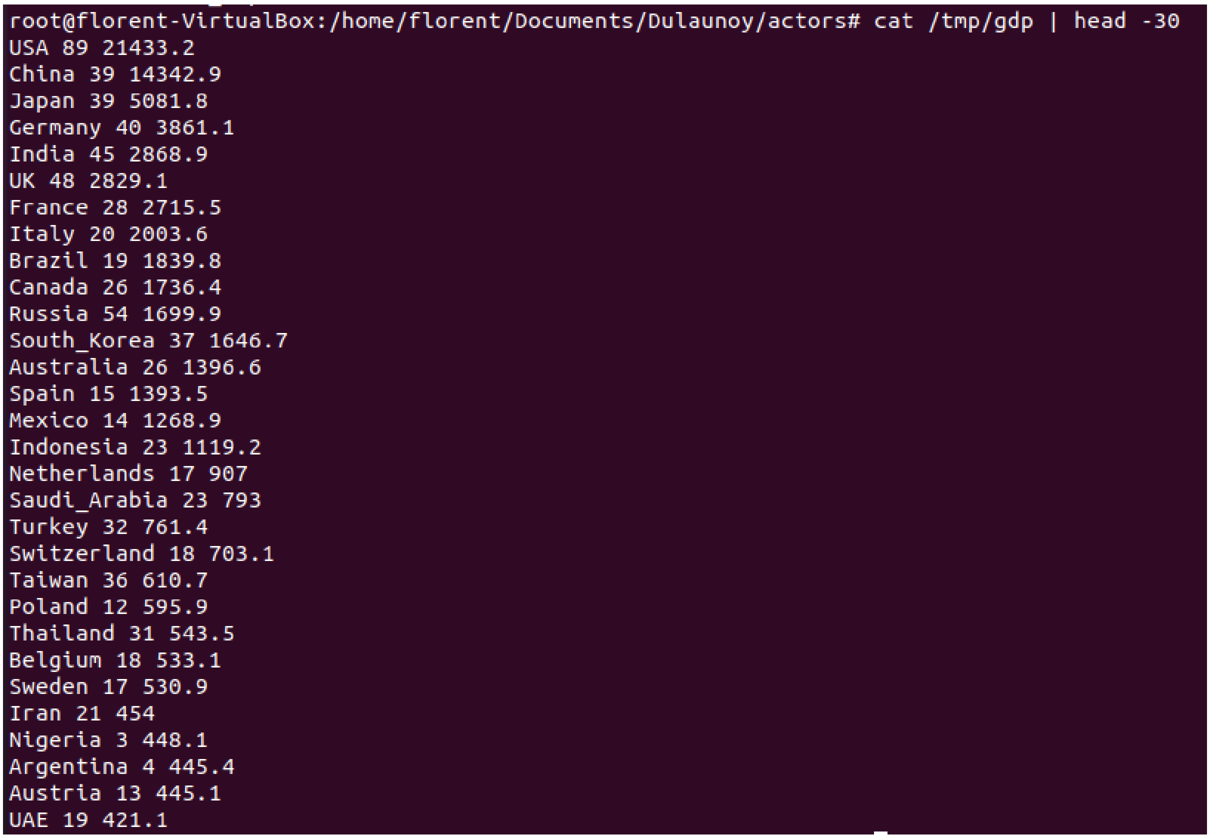
It doesn’t tell us anything, except that the USA and China “seem” to be less targeted, compared to other countries.

Let’s see what happens if we plot for every countries:



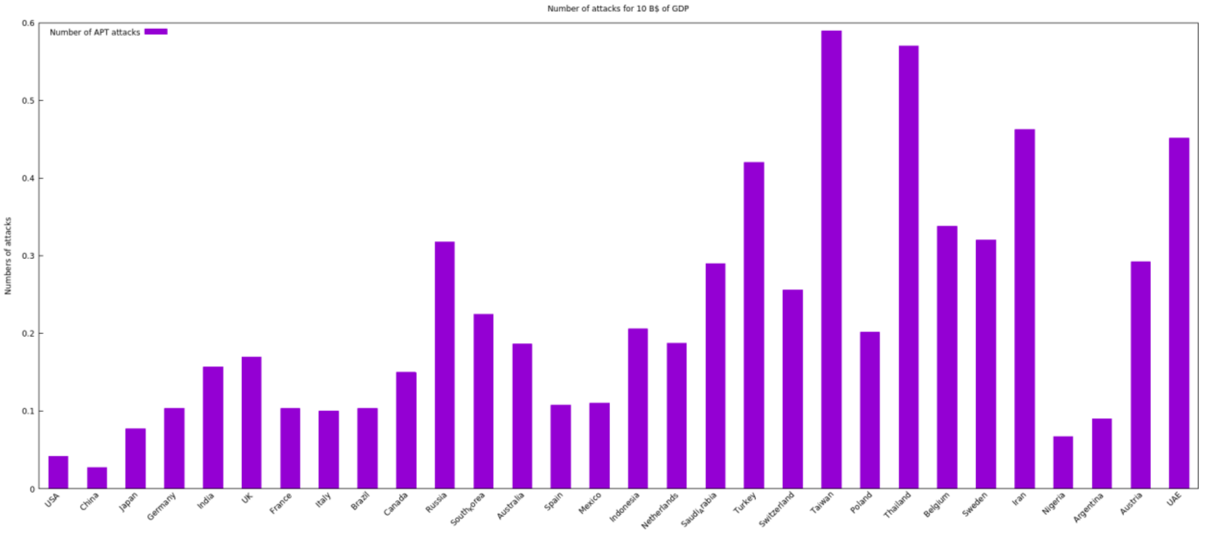
With this picture we notice that, by looking at proportion for 100M$ military budget, countries which have a small military budget are the most attacked, in proportion of course.

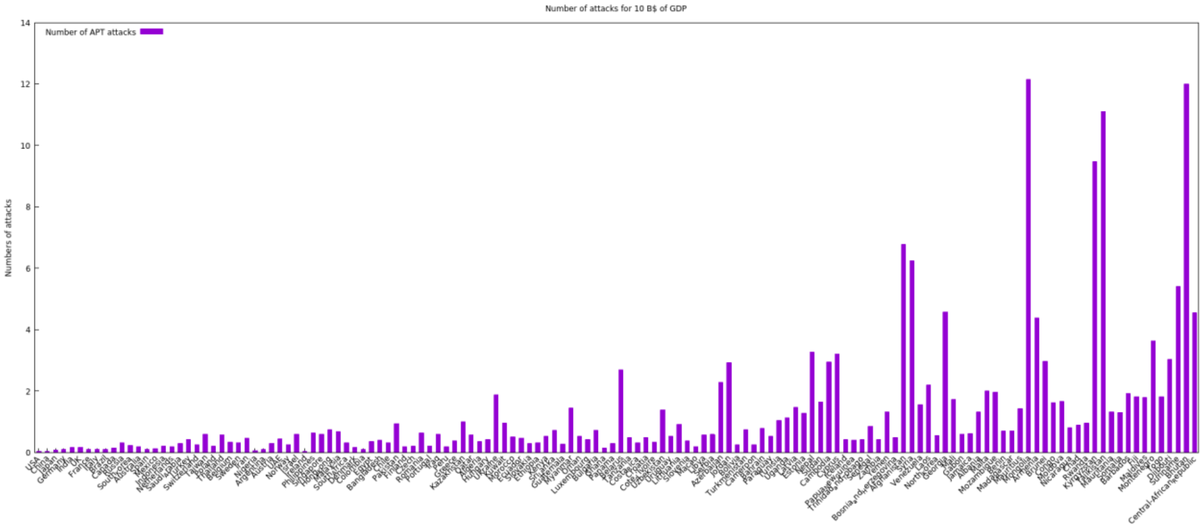
Let’s do the same thing concerning GDP field.



```
root@florent-VirtualBox:/home/florent/Documents/Dulaunoy/actors# cat /tmp/gdp | head -30
USA 89 21433.2
China 39 14342.9
Japan 39 5081.8
Germany 40 3861.1
India 45 2868.9
UK 48 2829.1
France 28 2715.5
Italy 20 2003.6
Brazil 19 1839.8
Canada 26 1736.4
Russia 54 1699.9
South_Korea 37 1646.7
Australia 26 1396.6
Spain 15 1393.5
Mexico 14 1268.9
Indonesia 23 1119.2
Netherlands 17 907
Saudi_Arabia 23 793
Turkey 32 761.4
Switzerland 18 703.1
Taiwan 36 610.7
Poland 12 595.9
Thailand 31 543.5
Belgium 18 533.1
Sweden 17 530.9
Iran 21 454
Nigeria 3 448.1
Argentina 4 445.4
Austria 13 445.1
UAE 19 421.1
```

We’ll plot the number of attacks towards a country, for 10 B$ of GDP:





Once again, we can see that the poorest countries are the ones who are more attacked, if we talk about proportionality.

Conclusion and improvements

With this project, we manage to discover data about APT groups, MISP galaxies, and we refreshed our bash knowledge to script.

The script is quite static since we have to modify it to add some fields from the Country_list file. It would have been better if it asked for the fields we wanted to deal with, then export those fields with APT group, or country respect.

We could have gone deeper in the granularity for regimes and region. Indeed, if we do that, we can spot that some APT groups target some specific regions, like did Roaming Tiger, which mainly target ancient USSR Republics.