Replication "David Versus Goliath: Fundamental Patterns and Predictions in Modern Wars and Terrorist Campaigns"

Last update: 2018 06 22

This repository contains all the material needed to replicate the results in the paper *"David Versus Goliath: Fundamental Patterns and Predictions in Modern Wars and Terrorist Campaigns"*. The analysis was carried out in R, and information on the session can be consulted in `session_info.md`. There are three folders

1. `data`
2. `code`
3. `output`

whose contents are described in more detail below.

If there are any issues/difficulties in replicating the results feel free to contact Stijn van Weezel: stijn.vanweezel[at]ucd.ie.

data

Containt the original data used for the analysis. Concerning the terrorism data from the GTD we included the RData file we created in order to save space as the original excel file had a size of about 80Mb. Note that rather than replicating the results, there is also the option of conducting a re-analysis using updated data from the provided data sources: both UCDP and GTD update their data normally in June.

- **ged171.Rdata**, conflict event data from UCDP (note that this is actually version 17.2)
- **globalterrorismdb_0617dist.RData**, data on terrorist attacks from GTD

code

This folder includes all the code used to i) process the data, ii) run the analysis, and iii) create the figures. First step is to fit a power-law to the data, using the following code

1. `prepare_data.R`, cleans the conflict event data and should be run first
2. `fit_power_law.R`, fits a power law to data and estimates the alpha parameters along with the xmin value
3. `fig_results.R` plots the estimated alpha parameters along with their p-value (figure 2 in paper)

Next step is to cross-validate the results, this is done by using the following code

1. `prepare_cv_data.R`, prepares the data for the cross-validation exercise
2. `cross_validation.R`, runs the cross validation exercise (produces figure 4 and 5)

A number of extra tests are executed. First, a power law model is fitted to data on terrorist attacks, which is done using

1. `prepare_data_terrorism.R`, which cleans the data on terrorist attacks
2. `fit_power_law_terrorism.R` fits a power law model to the data and plots the results (figure 3 in paper)

Second, for the conflict event data a log-normal model is fitted 1. `fit_logn.R` fits the log-normal model 2. `cross_validation_logn.R` carries out the cross-validation exercise and is used in combination with `cross_validation.R` to plot the results (figure 6)

The example in figure 1 is produced by `fig_eg_power_law.R`.

output

Contains the results as produced by the scripts in the *code* folder.

1. `fit_power_law.R` will create `fitted_values.csv`, `results.RData`
2. `cross_validation.R` will create `rmse.pl.RData`
3. `fit_logn.R` will create `results_ln.RData`
4. `fit_power_law_terrorism.R` will create `fitted_values_terrorism.csv`, `results_terrorism.RData`