Top 100 historical figures of Wikipedia

The **top 100 historical figures of Wikipedia** were determined by researchers from the University of Toulouse in France using mathematical and statistical methods from the Wikipedia database, and published in two scientific papers. In the statistical respects this top 100 list is of differs from the historical, cultural and other type arguments used by such historians like Michael H. Hart. The various mathematical methods and results obtained by different groups are described below. In spite or the mathematical and statistical grounds of those approaches they have overlap of about 43 percent with the top 100 list of Hart. The distribution of top PageRank historical figures over world countries is shown in Fig.1.

Approaches of different groups

The early ranking of top people of Wikipedia was done on the basis of PageRank algorithm and HITS algorithm for English Wikipedia edition (2005) by F.Belloni and R.Bonato .[1] For top people of PageRank they found Jesus, Paul the Apostole, Saint Peter and for HITS George W. Bush, Adolf Hitler, Bill Clinton.
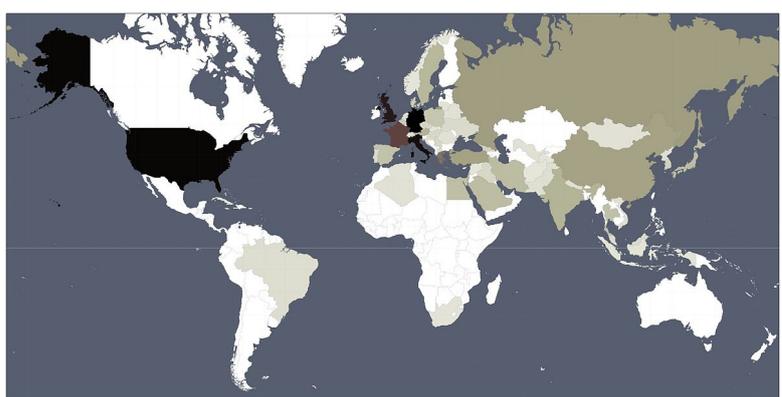

Fig1. World map of top 100 historical figures of 24 Wikipedia editions from PageRank for the period from BC 15th century to AD 20th century (darkness is proportional to a number of appearances of top persons born in a given country of AD 21st century geographical borders). After.

Later studies of Quantware group analyzed English Wikipedia edition Aug 2009 using PageRank, CheiRank and 2DRank algorithms.[2] The top persons found are: Napoleon, George W. Bush, Elizabeth II for PageRank; Michael Jackson, Frank Lloyd Wright, David Bowie for 2DRank; Kasey S. Pipes, Roger Calmel, Yury G. Chernavsky for CheiRank. For this study the distributions of top 100 historical figures of PageRank, CheiRank and Hart's list are shown in Fig.2.

Time evolution of Wikipedia ranking of historical figures was investigated for English Wikipedia editions for 2003 - 2011 using the approach developed for Wikipedia Aug 2009.[3] The distribution over fields of human activity was established there for various years.

Independently, a study with 15 largest Wikipedia language editions was done by Barcelona Media group .[4] This group considered network of links between biographical articles of Wikipedia. However, a number of such biographical articles is relatively small compared to the total number of articles of a given edition that led to fluctuating ranking results.

The investigations of 9 Wikipedia editions have been reported by Eom and Shepelyansky producing a reliable ranking of top 30 persons for each edition. However, a selection of historical figures from the whole list of ranked edition articles was done manually that was restricting efficiency of the approach.

In parallel, the Stony-Brook group performed ranking of English Wikipedia edition combining
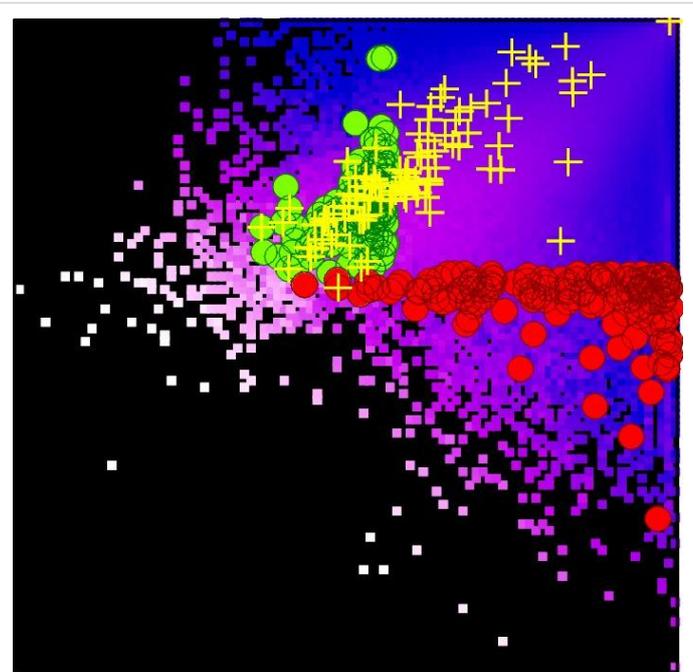

Fig2. Density distribution of Wikipedia English articles (Aug 2009) in the plane of PageRank and CheiRank indexes $0 < \ln K, \ln K^* < \ln N$ shown by color with blue for minimum and white for maximum (black for zero); green/red points show top 100 personalities from PageRank/CheiRank, yellow pluses show top 100 personalities from Hart's book, number of articles $N = 3282257$. After.

PageRank method with other methods .[5] This group found the top figures: Jesus, Napoleon. Muhammad. However, even if this group used the public Wikipedia database the whole list of their top 100 people is not publicly available.

The Pantheon MIT project produced the ranking list of top 100 persons using all language editions of Wikipedia counting number of editions and clicks on an article about a given person.[6] This group found at the top: Aristotle, Plato, Jesus.

A list of the top 100 historical figures was created from Wikipedia pages in 24 different languages, using computer algorithms to analyze the importance of people based on the links to those people's pages.[][][7]

Ranking Methodology

The researchers used several different page-ranking algorithms, including Google PageRank, 2DRank, and CheiRank. They retrieved data from the text of Wikipedia pages in the 24 languages, and applied the algorithms to the data to create culturally-specific list of influential people, as well as a list across all the cultures examined in the project.

Among the data elements specifically targeted as indicators of importance were each person's birth country, date of birth, century of birth, and quantity of hyperlinks. In the case of hyperlinks for people's Wikipedia pages, both links to a person's page and links from that person's page were included in the analysis.

Other methods are described at and, they are not directly related to link analysis, network theory and Markov chains.

Results

For the global list of 24 editions of Wikipedia, the top 10 historical figures, identified by averaging over PageRank lists, were as follows:

1. Carl Linnaeus
2. Jesus
3. Aristotle
4. Napoleon
5. Hitler
6. Julius Caesar
7. Plato
8. William Shakespeare
9. Albert Einstein
10. Elizabeth II

The top global persons of 2DRank are Adolf Hitler, Michael Jackson, Madonna (entertainer). The top women of human history are Elizabeth II, Mary (mother of Jesus), Queen Victoria for PageRank list and Madonna (entertainer), Elizabeth II, Mary (mother of Jesus) for 2DRank list. Top 100 historical figures for 24 Wikipedia editions are available at.[8] The overlap of top 100 people of Quantware, Stony-Brook, MIT Pantheon groups with the Hart list is found to be on a level of 42-44 percents. This shows that the mathematical methods of determination of top 100 historical figures of humanity via Wikipedia database give the relable results.

Discussion of Wikipedia ranking of historical figures in public press can be found at.

References

[1] Belloni F and Bonato R, "Network analysis for Wikipedia . Proceedings of Wikimania 2005, The First International Wikimedia Conference. Frankfurt, Germany. Retrieved September 16, 2006" (http://www.fran.it/blog/2005/08/network-analisis-for-wikipedia.html)

[2] Zhirov A O, Zhirov O V and Shepelyansky D L, "Two-dimensional ranking of Wikipedia articles, Eur. J. Phys. B 77: 523 (2010); DOI: 10.1140/epjb/e2010-10500-7" (http://www.quantware.ups-tlse.fr/QWLIB/2drankwikipedia/)

[3] Eom Y-H, Frahm K M, Benczur A, and Shepelyansky D L, "Time evolution of Wikipedia network ranking, Eur. Phys. J. B. 86: 492 (2013); DOI: 10.1140/epjb/e2013-40432-5" (http://www.quantware.ups-tlse.fr/QWLIB/wikirankevolution/)

[4] Aragon P, Kaltenbrunner A, Laniado D, and Volkovich Y, "Biographical Social Networks on Wikipedia - A cross-cultural study of links that made history, Proceedings of WikiSym, 2012 ArXiv:1204.3799(cs.SI) (2012)" (http://arxiv.org/abs/1204.3799)

[5] Skiena S, and Charles W, "Who's Bigger?: Where Historical Figures Really Rank. Cambridge University Press. ISBN 978-110704137" (http://www.maa.org/publications/maa-reviews/whos-bigger-where-historical-figures-really-rank)

[6] The Pantheon MIT project "The Pantheon MIT project" (http://pantheon.media.mit.edu)

[7] "Wikipedia Mining Algorithm Reveals The Most Influential People In 35 Centuries Of Human History" (https://medium.com/the-physics-arxiv-blog/wikipedia-mining-algorithm-reveals-the-most-influential-people-in-35-centuries-of-human-history-ede5ef827b76)

[8] Top people of Wikipedia "Top people of Wikipedia database" (http://www.quantware.ups-tlse.fr/QWLIB/topwikipeople/)

External links

- The Anatomy of a Search Engine, describes the PageRank algorithm developed by Google founders Sergey Brin and Lawrence Page (http://infolab.stanford.edu/~backrub/google.html)
- Dima Shepelyansky, Laboratoire de Physique Théorique, Université Paul Sabatier, France (http://www.quantware.ups-tlse.fr/dima)

Article Sources and Contributors

Top 100 historical figures of Wikipedia *Source*: http://en.wikipedia.org/w/index.php?oldid=616877760 *Contributors*: Animalparty, Frmorrison, HarryBoston, OccultZone, Roscelese, Shepelyansky, Spinningspark, Topbanana

Image Sources, Licenses and Contributors

Image:PageRank24x100.jpg *Source*: http://en.wikipedia.org/w/index.php?title=File:PageRank24x100.jpg *License*: Creative Commons Attribution-Sharealike 3.0 *Contributors*: Sfan00 IMG, Shepelyansky

Image:CheiRank3.jpg *Source*: http://en.wikipedia.org/w/index.php?title=File:CheiRank3.jpg *License*: Public Domain *Contributors*: Shepelyansky

License