Web Based Reputation Index of Turkish Universities

Sadi Evren SEKER

Outline

Question

- What do you think? Is there a correlation between the reputation and Internet web activities of a university?

- How would you model the correlation?

Introduction

University Web Reputation Index

Min-Max Normalization

Population Normalization

Web Mining

University Web Index

Number of Backlinks

Page Rank

Internet

Page views per user

Facebook Likes

Parameters

♦ **Has a Facebook Page?** We consider the social media as a part of the web reputation and we have checked if the university has a Facebook fan page or not. Only 128 of the universities among 170 have a Facebook page.

♦ **Facebook Like Count**. Another indicator is the number of likes of the Facebook fan page. Among the 128 universities who has a Facebook page, the maximum number of likes is 71114 and the average like count is 7817.

♦ **Value of the Site**. Some of the independent organizations offers a free agent to calculate the expected value of the web site via the web indicators like Alexa ranking or Google page rank. Most of them are built on the number of visitors and expected click from the visitors to make a valuing. The maximum expected value of a university web page in Turkey is 326.642 and average value of the university sites in Turkey is 11.965 USD.

- **Yahoo! Backlinks**. The number of backlinks is provided by Yahoo!. Depending on the crawler of the Yahoo!, the backlinks are counted by the number of other web pages holding links to the university web page. The maximum number of Yahoo! backlinks is 80400 and the average number of back links is 7749.

- **Google Backlinks**. The number of different web pages holding a link to the university web page depending on the Google web crawler. We need to indicate that the number of backlinks counted by Google is much more higher than the yahoo crawler. This might happen because of deeper crawling by Google. The maximum number of back links is 84 million while the average is 17million per university web page.

- **DMOZ Index** is a open directory project where the web pages are listed via their categories. We have checked if the web pages of the universities are listed in DMOZ (Directory MOZilla). This indicator has no effect on the result because all of the universities have a DMOZ listing.

Parameters

- **Number of Google Indexed Pages.** Another indicator is the number of web pages indexed by Google. The web pages of the universities have different number of subpages. For example the maximum number of Google indexed page is about 9 million while the average number of Google indexed page is 234.883. The average number of Google indexed page indicates that an average university may hold about 234.883 web pages in the same domain which may be the home pages of the professors, the student pages, the course pages, the administrative pages or the announcements.

- **Yahoo Indexed Pages** is another indicator in our indexing study. The Yahoo indexed page count is relatively less than the Google count, where the maximum is about 4.400.000 and the average is 34.327.

- **Daily Unique Visitors** is the average number of visitors per day. Also this number is seasonal and can change from season to season. In order to make a clean room measurement, we have collected the information during the off season, which is the summer term and most of the universities has no attendants. As expected the number of unique visitors is increasing during the final and registration weeks and since all the universities have different calendar we have collected the information during the off-season. The maximum visitor is 19.718 and the average is 757.

Parameters (Cont.)

- **Plagiarism** is another web indicator, where there are online web robots trying to find similarities on the web site content and other resources. These similarities cannot be considered as academic plagiarism in most of the time. For example a paper published by a professor is both published in his personal home page on the university web page and another academic directory at the same time. These robots can consider this type of duplicate publishing as a plagiarism. Although the plagiarism robots are not good in finding an academic plagiarism in most of the time, they are quite good in indicating the number of fresh publishing from the university academic staff. We have included the number of plagiarisms returned for each of the university web page as another indicator in our study and the maximum value is 10 while the average is 4.

- **Speed Test**. We have also executed some speed tests from 5 different global locations. The speed tests are simply executed by the ping rates and the average of 5 ping rate is normalized with a higher weight to the speed test from Turkey since most of the visitors are connecting those sites from Turkey. The ping rates are also gathered in different days times in order to avoid a temporary site failure. The maximum ping rate in average calculated is 3208 miliseconds and average is 240 ms.

- **Alexa Ranking** is another indicator published by an Amazon owned web site alexa.com. The lesser number means the web page has a higher ranking and the minimum ranking for the Turkish university web pages is 924 and highest ranking is 26992405 among the whole web sites on the Internet. The rankings above 20 million can be considered as a fresh web site.

Parameters (Cont.)

- **Alexa Bounce Rating** is another rating collected from alexa.com and indicates the number of bounces which means the users just visit a single page and then leaves the web site. The higher rates lower the web reputation index while the lower values indicate a higher reputation value. The maximum value is 90% and minimum is about 7%.

- **Page Views Per User** is another indicator to calculate the number of pages visited by a single user. The higher number means the user is spending more time to visit more pages and we consider this time spending as an indicator to a more attractive web site. The maximum is 6.40 and the average is 1.83.

- **Time on Site** is a web indicator to measure the time spending of the users with a time interval of their entrance and exit. The higher time means a higher reputation for the web site and the maximum value of time spent on the web page is about 13 minutes and average is about 4 minutes. These time intervals are also daily, which means the time on site indicator is an average day based time spending on web page for each of the user.

Parameters (Cont.)

- **The number of Sites Linking In** is another indicator that the number of inner links is counted. The number of inner links can be reviewed from two different perspectives. The former view can be the bigger web sites, which are relatively holding higher reputation has a high number of inner links. The latter view can be the strongly connected web sites where the users can navigate from one page to other easily and therefore the reputation is higher again, has higher inner link count. The average is 908 and maximum is 9200 for this indicator.

Normalization

- Min-Max Normalization

$$N_{Min\ Max}(x) = \frac{x - Min}{Max - Min}$$

- Why? For example the number of Google backlinks is varying from 80 million to 2 million while the speed test is varying from 3.2 seconds to 240 mili seconds. In order to get a final value from all these indicators we need a common range. The min-max normalization value always results between 0 and 1.

Index Creation

- Web Based Reputation Index

- Population Normalization

$$WRI = \frac{\sum_0^C N_x - \sum_C^K N_x}{C}$$

$$N_{pop} = \frac{WRI}{P}, \begin{cases} P = 1, if\ population = 0 \\ P = popluation, if\ population > 0 \end{cases}$$

Data Set

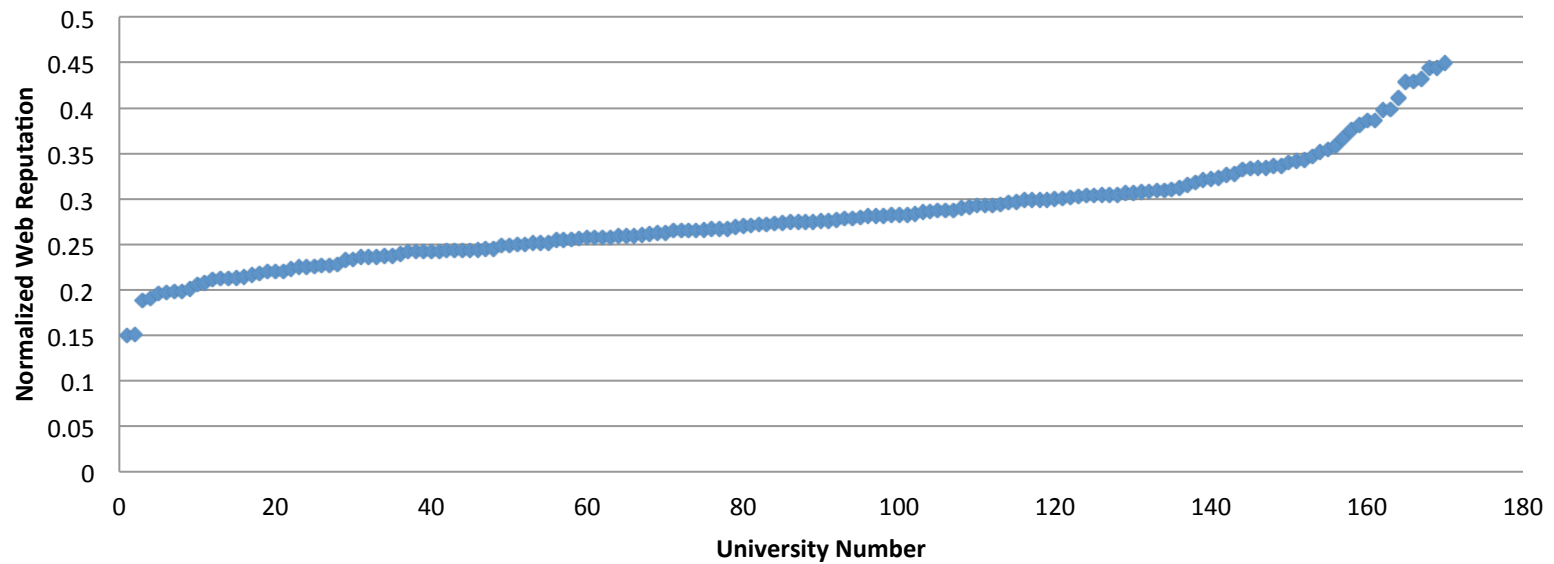TABLE I: PROPERTIES OF THE INDEX VALUES

Mean (μ)	0.280
Maximum	0.449
Minimum	0.150
Standard Deviation (σ)	0.055
Total Number of Universities	170

Visualization of Index

Distribution of Web Reputation Index

Top 10

TABLE II: THE MOST REPUTABLE TURKISH UNIVERSITIES ON THE WEB

University	Value
Anadolu	0,449508
İstanbul	0,444084
Gazi	0,443798
Bilkent	0,431946
Sakarya	0,42942
Boğaziçi	0,428301
Hacettepe	0,411034
Ankara	0,398226
Yıldız Teknik	0,397495
Dokuz Eylül	0,386229

Lowest 10

TABLE III: THE LEAST REPUTABLE TURKISH UNIVERSITIES ON THE WEB

University	Value
Deniz Harp Okulu	0,150473
Kara Harp Okulu	0,151046
Karatay	0,188876
Ankara Bilge	0,19034
Karabük	0,196542
Tunceli	0,197094
Şırnak	0,197622
K. Mehmetbey	0,198129
Avrasya	0,200475
Ağrı İ. Çeçen	0,205266

Conclusion

- Web reputation can be completely different than the rest of university ranking systems.

- For example, military academies (which are not the worst universities and relatively they are with up-mid reputation at least) are ranked lowest. Or a mid-tier university with a huge number of distant education system is ranked as top-university

- But, E-Reputation has an increasing trend and university rankings will be effected/effecting sooner or later.

Thank You

Sadi Evren SEKER

www.SadiEvrenSEKER.com