Journal of Web Engineering, Vol. 13, No.1&2 (2014) 114-128

EMPIRICAL STUDY OF LOAD TIME FACTOR IN SEARCH ENGINE RANKING

JAKUB MARSZAŁKOWSKI JĘDRZEJ M. MARSZAŁKOWSKI MACIEJ DROZDOWSKI

Poznań University of Technology

jakub.marszalkowski@cs.put.poznan.pl

jedrzej.marszalkowski@cs.put.poznan.pl
maciej.drozdowski@cs.put.poznan.pl

Received November 12, 2012
Revised December 5, 2013

Search engine ranking position is essential for marketing plans of many e-businesses. There is a lot of confusion over factors influencing this position. One of them is time performance factor. The purpose of this paper is to identify the role of load time factor in the algorithm ranking the search results in Google. We present empirical study of 40 phrases and load times measured for the first 30 results for each phrase (1200 websites). Two types of load time factors used by Google were analyzed. To quantitatively confirm the results, simulation of the algorithm is proposed. The simulations show that the load time factor plays role in the Google algorithm, although its weight is not high. From two studied metrics Google seems to use the crawl time, i.e. time spent downloading a page by Google robot. A second set of experiments confirms that load time is not only correlated with the ranking position, but also effectively determines this position.

Key words: E-commerce, Search Engines, Load time, Performance, SERP
Communicated by: M. Gaedke & N. Koch

1 Introduction

According to comScore [7] in December 2009 internet users made over 131 billion search requests worldwide. Google served 66.8 percent of this huge market. A more detailed view comes in the survey of local markets and Google shares [5]. Google global share is lowered by a few large countries like China and Russia with dominating local search engines (Baidu and Yandex), and by a few other large markets like US or Japan with considerable share of Bing or Yahoo! In other countries more than 90% searches are served by Google. Search engines, especially Google, are crucial for marketing plans of many e-businesses, because they control the flow of the network traffic. As shown in IAB report [24], e-commerce companies spend 16.9 billion dollars yearly to buy network traffic by search engines advertising. Traffic is obtained both from sponsored and organic results. Though Sen [36] argued that SEO is never an optimal marketing strategy, newer empirical studies show that both organic and SEO approaches are important, and influence each other [42]. Serano [39] showed the practical importance of being well positioned in e-commerce. Online sales can be built on sponsored search advertising which received a lot of attention recently [25, 34, 35, 40, 43]. Network traffic obtained by gaining a high position on Search Engine Results Pages (SERPs) is also very important for marketing strategies of many websites [13, 23]. However, research on the determinants of SERP positions and factors is limited.

Load time, the time between sending a request for an object, and receiving the object, plays an important role in e-commerce. Research by Gutzman [23] showed in 2000 that there is an eight seconds rule, i.e. if the page loads longer than 8 seconds, then there is a good chance that a customer will direct his/her browser elsewhere. In 2006, this eight seconds rule has halved [6]. To the best of our knowledge, there are no newer elaborations on this topic, but as internet connections are getting faster, it can be expected that this time is getting shorter. According to the Amazon experiments, every 100ms delay in page load time decreases the sales by 1 percent [27]. Similar measurements at Google showed that 500ms delay in search result display reduces revenue by 20 percent [4].

Thus, SERP rank and load time are two important factors influencing e-commerce profitability. It can be expected that they are mutually related. In this paper we study the influence of web page time performance on its rank in SERP. Due to the leading role on the market our research will focus on the case of Google, with the exception of Section 5, where other search engines will be included. Two different time factors will be analyzed. An experiment will be performed to validate several claims made by Google, and answer open questions posed by commentators. The main contributions of this paper are as follows:

- a new approach to web page load time beyond the point of view of customers psychology;
- a survey and a new contribution to the methodology of analyzing search engines algorithms, and more generally unknown ranking algorithms;
- validation/falsification of some Google claims on its algorithms, as well as some popular speculations about it (cf. section 2.1 for list of claims and questions);
- quantitative estimation of the weight of the load time factor;
- some additional insights into Google and other search engine ranking algorithms gathered by experiment with artificial websites.

The rest of this paper is organized as follows. In the next section research on SERP position factors is surveyed. In section three, the load time factor is analyzed. The experiment performed to find the role of the load time factor is described in the following section. The experiment with artificial websites and its results are presented in the Section 5. Section 6 outlines the results and discusses them. The last section comprises conclusions.

2 Known factors and previous research

There are many factors determining the SERP position and thus creating the whole Search Engine Optimization (SEO) substantial market. SEO found its way to the university level education [12]. There are many books meant to guide into the art of SEO. Examples include [9, 10, 21]. These guidebooks list numerous factors influencing position in SERPs and offer advice on optimizing the pages. The factors are usually divided into on-site and off-site. On-site factors apply to the content (i.e. the phrases), its importance (HTML emphasis like bold or headers), its readability (HTML structure), etc. Off-site factors apply to page popularity over the internet (links to the page, quantity, quality and phrases used in the links). A group of query independent factors (trust, social response, etc.), which are often far beyond technology, exists as well. Many SEO practices can be considered unethical. Google itself does not forbid SEO. They rather discourage so-called “black SEO”, and provide guidelines to make pages more readable both for people and robots [20]. A popular rule for the SEO ethics says that on-site SEO is acceptable ("white hat"), while the off-site is denied ("black hat"). Yet, this view is probably too simplistic and does not cover the real-world complexity. Jones [26] points that unethical SEO includes use of on-site spam and hiding techniques. He also outlines these techniques and provides analysis of numerous papers on fighting them.

It should be noted that SEO not always has rigorous scientific justification. Most of the factors are just publicly believed to play a certain role. Confirming their importance, or doing any reverse engineering of the Google algorithm, is problematic due to the following difficulties. Google keeps the factors and the SERP ranking algorithm secret. They are of utmost value for SEO industry, as even Google admits that if the algorithm was known, then it could be outplayed [41]. Thus, the algorithm was a subject of several studies. Sedigh and Roudaki [37] used linear regression analysis of several observed parameters to calculate ranks and came to the conclusion that accurate identification of the Google's algorithm is not possible. Bifet et al. [3] used support vector machines and binary classification trees to analyze impact of over 20 factors, but achieved only a limited success. Evans [11] took data from the results of a SEO competition and analyzed (mostly visually) 8 factors that are query independent. As new factors emerge, they happen to be analyzed too. Kowalski and Król [28] experimented with web traffic factor. In March 2011 Google introduced a new algorithm codenamed Panda [41], to get rid of pages that were outplaying the old algorithm. New factors introduced by Google are supposed to reflect: content uniqueness, its value for the users, trust in content creator (usually the site), page usage by the users, social response both to the content and the entire site. The annual report of SEOmoz company analyzes the post-Panda status [38]. They not only surveyed 134 SEO industry professionals for their opinion on the future shift in factors importance, but also analyzed 10,271 keyword search results. Numerous factors were analyzed using Spearman rank correlation coefficient. They measured one kind of response time once and concluded that it was not playing any role. This might be a result of focusing their research on the factors related to page content, rather than the load time.

2.1 Load time factor

For a long time site speed has been rumored as playing some role in Google's algorithms, though there was no evidence for this. Every reputed set of SEO tools offered utilities for making diversified speed measurements. Google was always paying attention to websites load speeds. They conducted their own research proving that speed does matter [4]. They also provide a set of articles, lectures and talks on making websites work faster [15]. Finally, in April 2010 Google announced site speed as a new factor in their ranking algorithm [17]. Through this and several other channels [8], Google claims that:

- relevance is still the most important factor,
- fewer than 1% of search queries are affected by the load time factor,
- it affects mostly slow sites.

Meanwhile in the commentaries to the inclusion of the speed factor into the ranking algorithm [17], on SEO blogs and e-zines [31, 32], many uncertainties and open questions were posed:

- Which of the two metrics (see section "Two speed metrics") that Google gathers, is used?
- What kind of queries fall into the affected 1%?
- What does "slow" mean for Google?

- How to understand the above claims?

Furthermore, there is no information whether the later updates changed the role of the load time factor. To the best of our knowledge, there was no comprehensive research on the load time factor influence on SERP positions. Moreover, taking care of the speed can never be an unethical SEO practice.

2.2 Two speed metrics

Google is gathering two metrics of site speed: crawl time and page load time.

The first metric (crawl time) is the file download time measured by the Google crawlers when indexing the web. This metric is accessible for the site owner in Webmasters Tools [14], in the section "Diagnostics/Crawl stats" described as "time spent downloading a page". There are min, max, average values as well as a chart for the last three months. This metric existed for a long time. SEO guides usually advice to keep this time reasonably low, not to slow down Google when indexing the website. Yet, no mention is given of its further role. Download times for several content types are included in crawl time [18]. However, the HTML files play a dominant role. At least three crawlers read them: Googlebot, Googlebot News and Google AdSense [19]. Other file types cannot have a great impact: files like PDF or Flash are the least numerous, CSS and JavaScript are very rarely read by crawlers. Only image files are fetched in larger numbers, but "site:" queries show that the number of indexed images per website is only around 10-50% of the number of indexed HTML pages. Images, and all the remaining file types are rarely reread to check changes, while it is done very frequently with HTML. Therefore, in the further considerations we take HTML downloading time as a representative for crawl time.

The second metric (page load time, a.k.a. page speed) is the time of loading and rendering a complete page with all the files required by HTML (JS, CSS, images) [16]. This metric is measured in user browsers by Google Toolbar with the PageRank feature enabled. It was accessible in Webmasters Tools, in the section "Labs/Site performance" since 2009 till November 2012. It showed the average time, a chart for the last six months, and information on the percentage of sites that are slower than the one currently analyzed. Information on the accuracy of these estimates was given as: low (fewer than 100 data points), medium (100 to 1000) or high (more than 1000). It should be noted that on many sites with noticeable traffic there was a "no data available" marker. Even for the sites with over ten thousand visitors daily, the accuracy still happened to be medium. There is no information whether after removing public access to the data in 2012, the measurements were stopped.

There are many hardly supported claims on the page load time metric [16]. For example, visitor locations, internet speed and behavior will be selected to represent real traffic of the site. Moreover, Google claims that 1.5 seconds is 20th percentile of the load time. It is marked as a fast/slow border on charts in Webmasters Tools [14]. The claim that a page with load time of over 1.5 second is slow may be Google equivalent of the 8 seconds rule for the current state of the Internet. The introduction of Analytics measurement added even more uncertainty around the role of the page load time [22]. The more this role deserves a study and clarification.

Recently page load time is also measured by Google Analytics script, and is available in this application with some more features. There were no official statements if this new measurement affects SERPs. However, data availability is even more limited here, as a website must not only have Google Analytics, but also a new measurement script to be benchmarked.

3 Experiment setting

The purpose of our experiment is to validate or disqualify the role of the load time factor in ordering the results on Google SERP. To be more precise, we want to verify the correlation between SERP position and the load time factor. As there are two factors Google could possibly use, we shall study if both or any of them is used. Another goal is to answer the questions raised in the previous section: whether only 1% of queries are affected by this factor, whether mostly slow sites are affected, and when a site is slow. We intend to verify if the load time factor works on the entire spectrum of phrases, or only a subset.

3.1 Research difficulties

There are many practical difficulties in studying Google ranking algorithm. Allegedly, there are over 200 factors, but there is no information on them or their weights [8,11]. The sheer number of factors illustrates the scale of the difficulty. Additionally, Evans [11] pointed that:

- the algorithm (the factors and their weights) can differ for different categories of phrases,
- Google has numerous datacenters with not always synchronized data - consequently returning different SERPs,
- the algorithm for the top set of results can differ from the one used for the remaining positions.

The last point should be clarified further. In the previous research it was assumed that Google's algorithm is continuous from the first to the last SERP position. Bifet et al. [3] analyzed top 100 results, Evans [11] top 50 or top 100 results, Sedigh and Roudaki [37] and SEOmoz [38] top 30. However, there are many indications that Google imposes penalties for unethical SEO, especially for breaking the Google "Guidelines for webmasters". For example, penalty named "minus 30" is supposedly moving a page 30 positions down, usually from positions 1-5 to positions 31-35 [2]. This has to create discontinuity. Possibly, it is not the only discontinuity, just the widely discussed one.

The selection of representative phrases is of great importance. Previous researchers analyzed from one phrase [37], a single phrase coming from SEO competition [11], through 48 phrases, with 28 selected for machine learning [3], to over 10,000 phrases [38]. Only the last study includes phrases from differing categories.

Another problem is whether the timings that the researchers register are the same as Google does. Obviously, time measurements can differ because they are made from different network locations or at different times. But the same happens for all other factors even those based on the page content, keywords, or HTML structure, because websites are more and more dynamically generated. Geolocation and web browser parameters can also affect the measurement. Time measurements variability can be partially eliminated by making large number of tests and averaging them. There seems to be no equivalent solution for the problem of variable page content.

3.2 Test phrases

Our experiment requires selection of the test phrases. They should not hinder observing the interesting phenomena, by restricting phrase lengths, categories of human activities, and most of all, different *popularity*. Phrase popularity is sometimes called difficulty or strength and is supposed to reflect how many different sites compete for the first SERP positions on the phrase. Phrase popularities are taken directly form Google results pages, from the row where Google claims the number of results they have for the searched phrase.

Table 1 Test phrases

set	no.	Phrase	popularity	set	no.	phrase	popularity
2	1	free classes	611 000 000	1	4	mortgage comparison	53 500 000
2	2	best credit card	536 000 000	2	18	online sports betting	52 100 000
2	3	top apple apps	468 000 000	2	19	electricity prices	43 700 000
2	4	last minute	423 000 000	2	20	injury claim	38 000 000
2	5	conference call	379 000 000	1	5	silver prices chart	30 500 000
2	6	engineering degree	323 000 000	1	6	diet plans for men	15 500 000
2	7	family attorney	282 000 000	1	7	wireless ip camera	10 900 000
1	1	car insurance	247 000 000	1	8	cyprus houses for sale	6 320 000
2	8	new smartphones	240 000 000	1	9	task management app	4 120 000
2	9	10” android tablets	202 000 000	1	10	one bowl chocolate cake recipe	2 710 000
2	10	sites that pay for blog posts	172 000 000	1	11	laser printer replacement parts	1 550 000
2	11	free web hosting	150 000 000	1	12	damascus steel knife	827 000
1	2	celebrity news	134 000 000	1	13	herbs for nail fungus	555 000
2	12	attorney general	127 000 000	1	14	fisheye lens effect online	312 000
2	13	donation center	109 000 000	1	15	cheapest hybrid suv	209 000
2	14	2012 olympics	98 000 000	1	16	running shoes for overpronators	146 000
2	15	forex account	85 400 000	1	17	rutabaga nutrition information	78 400
1	3	free fps games	75 000 000	1	18	bitcoin mining gpu	45 500
2	16	programmer job	74 300 000	1	19	blutwurst recipe	28 800
2	17	student loan	62 400 000	1	20	gambian pouched rats breeders	14 400

In the first run of the experiments we selected a set of 20 phrases shown in Table 1 as set 1, with ratio k=1.68 between consecutive popularities. Our preliminary results seemed to show some patterns in the subset of the five most popular phrases. To confirm this, we decided to select another set of 20 phrases, shown in Table 1 as set 2. These start with popularity above the fifth most popular phrase from set 1, and have the ratio of the consecutive popularities k=1.16. The relationship between the two sets is shown in Figure 1.

3.3 Measurement methods

In the experiment we had to collect time measurements similar to the two metrics gathered by Google. Both measurements were made using tools developed by Marszałkowski [29] and currently accessible as availability.pl [1]. For measuring crawl time, a PHP script with http_get function was executed on a server with several 100 Mbps parallel internet connections. Page load time was measured by a JavaScript component on remote terminal opening complete pages. In the experiment it was performed in Firefox browser on a 7Mbps internet connection. We measured two series of 20 test phrases. For each phrase we read results limited to the first 30 SERP positions, to avoid “minus 30” discontinuity described earlier. Each result is a web address – totaling 1200 web addresses. For each address we measured crawl times at least 400 times, and page load times at least 20 times. Over 500 000 data points were collected.

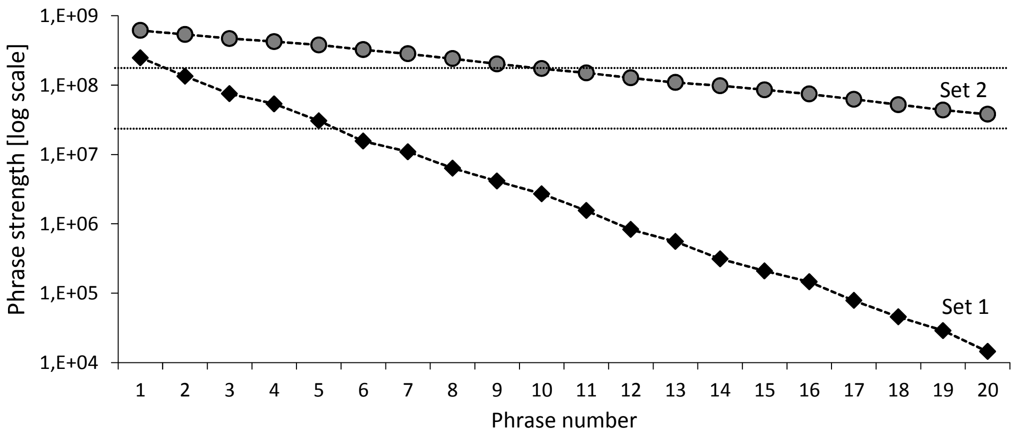

Figure 1 Popularities of test phrases.

4 Results

The collected data was analyzed with Pearson product-moment correlation coefficient calculated as [33]:

$$r = \frac{\sum_{i=1}^{n}(X_i - \overline{X})(Y_i - \overline{Y})}{\sqrt{\sum_{i=1}^{n}(X_i - \overline{X})^2}\sqrt{\sum_{i=1}^{n}(Y_i - \overline{Y})^2}} \tag{1}$$

where $X_i=i$ are SERP positions, Y_i are times measured for the address or the addresses on position i, $\overline{X}$ and $\overline{Y}$ are mean values and n is the number of samples. Here n=30 and $\overline{X}$ =15.5, because the first 30 addresses returned in the SERP were analyzed. As a rule, the closer r to 1 is, the stronger the correlation. Noticeable correlations were found by checking statistical significance of the results. This was tested by checking p-value by calculating

$$t_s = \frac{(n-2)r^2}{1-r^2} \tag{2}$$

and getting an F-statistic with 1 degree of freedom in the numerator and $n-2$ degrees of freedom in the denominator [30]. The p-value represents the chance that a given correlation was obtained randomly. We used the standard rule that p<=0.05 represents statistically justified value of r. The summary of observed values is presented in Table 2. We also tried Spearman rank correlation coefficient instead of Pearson correlation coefficient. It seemed to be better suited for this experiment: positions are clearly ranks of unknown distance between them and the dependence was probably non-linear. However, the results for Spearman coefficient were very close to the results achieved with Pearson coefficient, while the former was introducing additional difficulties. Of the 1200 tested addresses 10 were not answering the requests of the crawl time measuring script. This might be a result of some anti-crawler security settings. Spearman rank correlation coefficient could not be calculated for the phrases without complete time measurements, as there were ranks missing in the middle of the set.

As an indicator of the crawl time and the page load time for each web address we used medians taken from the entire set of measurements for each web address. This procedure removed the time measurement noise. There are 200 factors determining positions on SERPs. Their influence on page position may be understood as noise with respect to site speed performance ordering. To eliminate such noise we averaged the results over the SERPs positions for all phrases. Other time parameters were also studied: standard deviations, mean load times for the whole phrase set, its subsets divided by phrase popularities and for every separate phrase. Since they did not reveal any significant patterns we do not report on them.

4.1 Page load time

We compared page load time with Google claims about their data [16]. In our results the value of the 20th percentile for the complete set of times for all measured pages was 1.375 second. This value is close enough to 1.5 second claimed by Google, as this value is probably rounded and is 2 years old. For the complete set of data no correlation between SERP positions and the page load time was observed. Nominally r=0.07, p=0.698. This can be verified in Figure 2 in the right chart.

4.2 Crawl time

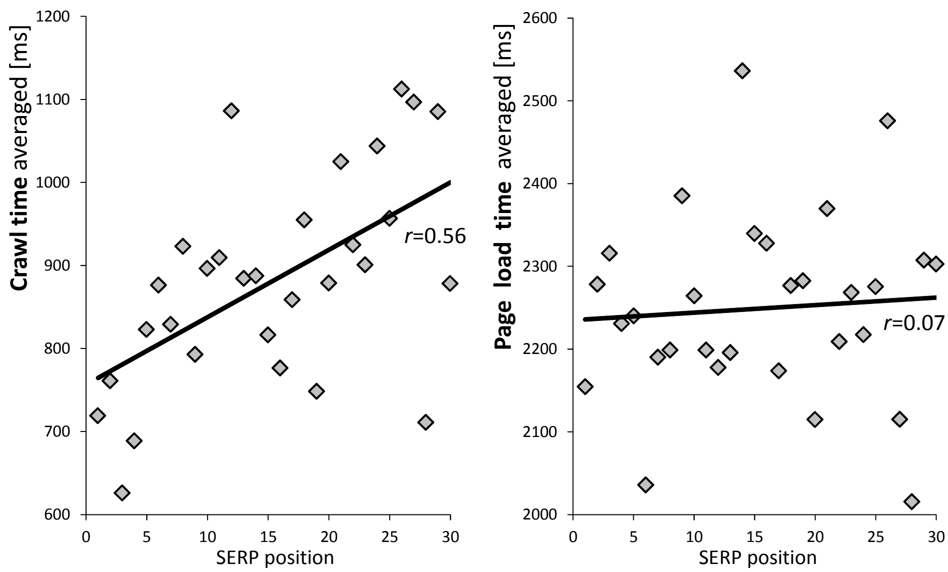

Figure 2 Times measured averaged for SERPs positions (in both charts Y axis has range of 600ms).

For the whole set of results, there is a significant correlation for crawl time, with r=0.56, p=0.001. This again can be visually verified in Figure 2 in the left chart. Similarly, the standard deviation of the crawl time is correlated with SERPs positions, as r=0.44, p=0.01. This is probably related, as faster sites are also more stable with respect to the crawl times. These correlations however are usually not visible in single phrases. Only in 7 of 40 phrases (17.5%) was the p-value below α=0.05 for the correlation between SERPs position and the crawl time (cf. Table 3). This is due to the alleged presence of the 200 other factors in Google algorithm. Moreover, in 70% of phrases the sites from positions 1-10 were on average faster than the sites from positions 21-30.

Table 2 Summary of discussed correlations (statistically significant are in bold)

Main experiment		
SERPs position correlations	*r*	*p*
crawl time	**0.56**	**0.001**
standard deviation of the crawl time	**0.44**	**0.015**
largest measured crawl time	0.36	0.053
90 percentile	0.24	0.199
crawl time 95 percentile	**0.37**	**0.044**
number of crawl times above 1.5s	0.33	0.074
page load time	0.07	0.698
Other tested correlations	*r*	*p*
phrase popularity with the load time	0.18	0.252

Artificial web pages experiment		
Load time correlations with SERP position	*R*	*p*
Google	**0.60 - 0.67**	**0.05 - 0.03**
Bing (early)	**0.73**	**0.01**
Bing (late)	0.26 - 0.38	0.41 - 0.22
Yahoo (pre Bing)	**0.64**	**0.03**
Baidu	-0.12 - 0.01	0.72 - 0.99
yandex	0.28 - -0.15	0.37 - 0.64

Table 3 Correlations observed on separate phrases (statistically significant are in bold)

phrase	*r*	*p*
electricity prices	**0.41**	**0.020**
rutabaga nutrition information	**0.40**	**0.023**
donation center	**0.39**	**0.026**
herbs for nail fungus	**0.39**	**0.026**
free fps games	**0.38**	**0.033**
celebrity news	**0.37**	**0.035**
top apple apps	**0.35**	**0.050**
blutwurst recipe	0.32	0.070
fisheye lens effect online	0.32	0.076
car insurance	0.29	0.111
forex account	0.28	0.117
diet plans for men	0.27	0.132
damascus steel knife	0.21	0.240
engineering degree	0.19	0.285
conference call	0.17	0.347
last minute	0.13	0.462
attorney general	0.13	0.479
mortgage comparison	0.13	0.486
laser printer replacement parts	0.12	0.502
freeclasses	0.12	0.529

phrase	*r*	*p*
injury claim	0.11	0.541
family attorney	0.09	0.641
silver prices chart	0.08	0.653
free web hosting	0.06	0.763
running shoes for overpronators	0.05	0.769
2012 olympics	0.03	0.868
10 android tablets	-0.05	0.774
online sports betting	-0.05	0.802
bitcoin mining gpu	-0.06	0.745
student loan	-0.06	0.752
one bowl chocolate cake recipe	-0.07	0.694
sites that pay for blog posts	-0.09	0.636
task management app	-0.10	0.581
cyprus houses for sale	-0.12	0.524
wireless ip camera	-0.13	0.492
new smartphones	-0.14	0.433
programmer job	-0.29	0.105
best credit card	-0.34	0.057
gambian pouched rats breeders	-0.40	0.024
cheapest hybrid suv	-0.43	0.015

4.3 Simulation

We used a simulation to:

- verify that the procedure of averaging measurements for a given SERP position over all phrases is working as noise removal method
- estimate quantitatively the weight of load time in the ranking process.

The value Y_i representing the analyzed factor was simulated as strictly correlated with position $Y_i=i$. The value X_{ij} imitates the *i*-th result of Google's scoring algorithm for phrase *j*. Following [8,11] we assume that 200 factors determine SERP. From the view point of the analyzed factor the remaining 199 factors are noise injected in the measurements of positions. To simulate it, random values in the same range from 1 to 30, representing noise coming from the 199 other factors were added to Y_i. This way we created the scores, and hence, the ranking for 30 positions. The position was strictly correlated with the simulated factor, but with limited weight w. Formally:

$$X_{ij} = wY_i + \sum_{a=1}^{199} \mathrm{RND}(1{,}30) \text{ for } i=1, \ldots, 30 \tag{3}$$

where RND(1,30) is a random number from range [1,30]. Simulated rankings were created for j=1, 2,..., 250 phrases. Then, we calculated averages over subsets of phrases, increasing the number of phrases included in the averaging:

$$X'_{ik} = \tfrac{1}{k}\sum_{j=1}^{k} X_{ik} \text{ , for } k=2,\ldots,250 \tag{4}$$

The correlation between X'_{ik} and Y_i was calculated as in the original experiment. For small values of k the results were of very random nature. Correlation coefficients were weak and below significance. With the increasing k, the correlation was becoming strong, and the p-values were showing statistical significance. As could be expected, the correlations between the averaged values and their positions were strong if the number of averaged phrases was high enough. On the other hand, for w=0 the X_{ij} was composed of uncorrelated factors and the correlations were approaching zero. This confirms that noise removal procedure works properly and the results are not of purely random nature. Several variations of the model in equation (3) were tested: Additive function with weights, discretized factor values (even binary), conditional inclusion of factors. As could be expected the results were alike. Beyond the above additive model, a multiplicative model was tested, giving the same results. Note that the case with weights is equivalent to the situation when some small groups of factors are correlated.

The simulation was run dozens of thousands times in an attempt to estimate the weight w of crawl time factor in Google algorithm, by matching the measured value r=0.44 with the results of simulated correlations. This was achieved by performing 500 runs for each tested w. We were changing the value of w to approach the real r value from above and below. For each tested w a set of j=40 simulated phrases was created, averages for positions X'_{ik}, for k=40 were calculated and then correlation between X'_{ik} and Y_i was calculated. The average weight found this way is 0.74% (w=1.5). By approaching the r with 95 percentile of the 500 experiments from above and below, we can make a claim that with 95% confidence level the weight of the crawl time factor lies between 0.35% and 1.04%.

4.4 Remaining findings

We observed no significant correlations between the phrase popularity and the load times. Nominally r=0.18, p=0.25. Such correlations appearing in the preliminary results, lead us to studying the second set of phrases. Yet, for the greater dataset the correlation between the phrase popularity and the load time factor was not reobserved. Clearly, correlations observed earlier were of random nature. However, it should be noted that there was a weak pattern showing that the sites in the popular phrases are a bit faster than the sites in the easier phrases. The difference in the average crawl time between the sites in the 10 most difficult phrases and the 10 least difficult ones was 12%. This is also visible in Figure 3.

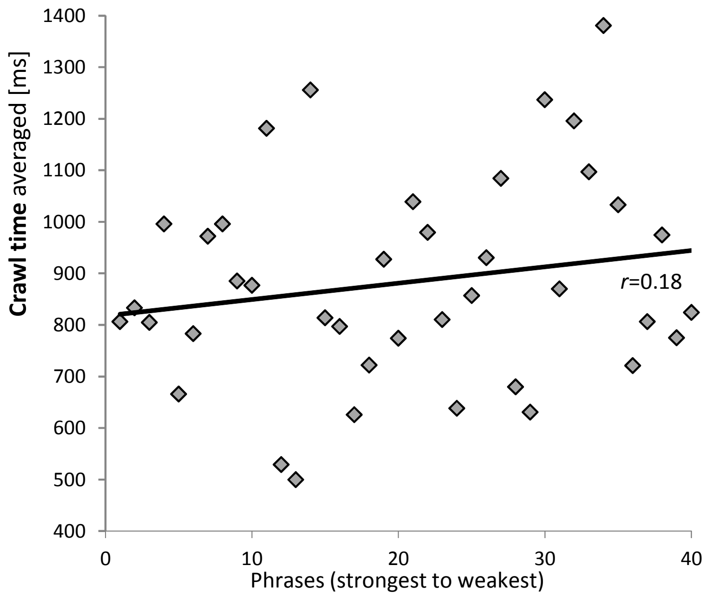

Figure 3 Crawl times vs phrase popularity.

No other pattern could be observed. The phrases having statistically significant correlation between crawl time and SERPs positions are highlighted with bold in Table 3. We were unable to find such pattern neither for the contextual categories of phrases, nor for the length of phrases, or any other. The results of section "Crawl time" suggest that all phrases together are affected by the crawl time factor, without preference for popularity. No evidence supporting the claim that "mostly slow sites are affected" was found. Firstly, slow sites are observed on all SERP positions even on the first one. This is true even with crawl time exceeding 1.5 second, which is believed to be slow for page load time. We also analyzed correlations of SERP positions paired with: the largest measured crawl time, 95 percentile, 90 percentile and the number of crawl times above 1.5s. As previously, these were calculated over all phrases. Only correlation for the 95 percentile was statistically significant with r=0.37 and p=0.044, but still weaker than r=0.56, p=0.001 observed for the medians. The other tests mentioned above had p-value above α=0.05.

5 Artificial web pages experiment

The correlation between certain page speed factors and SERP positions can be of diverse nature. There are at least two ways in which the above factors and SERP positions can be linked. On the one hand, the companies managing their pages better may appreciate both the importance of page time performance and of the other factors (e.g. quality of the content, usability). Then the remaining factors can promote the page to higher SERP positions, while the page speed is in no way determining the obtained rank. In this case, the correlation may arise as a side effect. On the other hand, search engines may give preference to the pages with high speed. Then page performance is not only correlated, but also determines SERP position. We intend to verify this in the final experiment.

We tailored 10 artificial websites on separate domains, servers and IP addresses. We created content for artificial phrase “GTX 615 Rainbowshade”. The websites were set to have response times from 0.15s to 3.45s with step of ca. 0.3s, incurred by applying sleep function in page PHP script. We used best of our SEO knowledge and practical experience, to construct websites in such a way, that factors other than crawl time would be the same and would not affect their position in SERP. The observations were conducted for 13 months. The load times accuracy was controlled in the same way as described in the Section 3.3, while positions in SERPs were analyzed as described in the Section 4.

Google: Two problems occurred with testing Google. Firstly, to get the web pages indexed in Google quickly, we linked them from a different web page of a high rank. To refrain the link position from influencing Google algorithm, we have put all links in random order in a single line. Unfortunately, this did not help because the second slowest artificial web page happened to be the first link. Consequently it received the first position in Google SERP. This page has been occupying the first position since then, while other pages positions were moving. This effect confirms a common knowledge that the value of a link pointing to a web page, in this case position in a list, is the factor of the greatest importance for the algorithm. Secondly, one of the web pages displayed an error page for a few weeks, due to its server failed update, which affected its position heavily. Both sites had to be removed from results leaving us with 8 sites. For these eight the correlation between crawl time and SERP position was between r=0.60, p=0.05 and r=0.67, p=0.03 for the last six months.

Initially we did not intend to extend this experiment to other search engines, but as time passed other search engines also indexed our web pages, and some interesting results occurred. As indexing happened organically the first problem mentioned with Google was avoided. Also most of other crawlers recheck web pages less often and did not notice failure of the second page. Thus all 10 pages could be analyzed.

Bing: Directly after inclusion of our pages in Bing, the results were strongly ordered by the crawl time, r=0.73 with p=0.01. However later on, the results were reordered due to some other factors and correlation became weak and below statistical significance. This might be a result of Bing having initially mostly this factor, and only latter gathering and applying some off-site factors.

Yahoo!: As long as we could observe Yahoo! Own results, there was quite strong correlation, r=0.64, p=0.03. Unfortunately currently all Yahoo! SERPs we can analyze are isomorphic with Bing SERPs.

Yandex: showed weak correlation then weak negative correlation, both below statistical significance.

Baidu: showed weak negative correlation, below statistical significance.

Other search engines both regional like seznam.cz, yahoo.jp or naver.com or alternative like duckduckgo.com index limited part of web, and included too few or even none of our web pages. Thus, they could not be analyzed.

A conclusion from this experiment is that load time determines SERP position especially when few other factors are available.

6 Conclusions

In this paper we analyzed the influence of the web page load time factor on the SERP position. A general conclusion is that the relation between the load time factor and SERP ranking exists, although it is not very apparent.

The metric of page load time (a.k.a. page speed) seems not to affect Google's algorithm generating the SERPs. This might seem reasonable as neither Google Toolbar nor Google Analytics are giving enough data for entire web. Furthermore, our research did not confirm any use of the page load time, in the range of popular phrases, where Google could have enough toolbar data to use. Instead, the crawl time measured by Google's indexing robot seems to be used. Results show that the load time factor affects all the queries and all the sites. The claim that "less than 1% of search queries are affected" might refer rather to the weight used for the load time factor in the entire ranking algorithm. The claim that "mostly slow sites are affected" cannot be supported. The impact of the load time factor in Google's algorithm is moderate, there are certainly more significant factors. The last set of experiments (Section 5) confirms that crawl time is not only correlated but also determines SERP position.

There is no information if the search engines other than Google are capable of gathering of reasonable amount of measurements for page load time metric. On the other hand, the metric of crawl time is very easily obtainable while indexing the web page. This can result in including crawl time factor in other search engines, as Bing and Yahoo! examples show.

It is not possible to grant a specific or even better SERP position with improvement of a single factor affecting the algorithm. Still, for many e-commerce projects improving load times can be a valuable move towards a better position. Web page performance is important also for better client experience. Thus, taking care of the load times should be advised for anyone involved in e-commerce.

Acknowledgements

This research was partially supported by a grant of National Science Centre NN519643340.

References

1. availability.pl. Availability.pl, 2012 (http://availability.pl, accessed Nov 8, 2012).
2. Bestbraindrain. The google minus 30 penalty, 2007 (http://bestbraindrain.com/2007/07/12/the-google-minus-30-penalty.aspx, accessed Jul 31, 2013).
3. Bifet, A., Castillo, C., Chirita, P.A., and Weber, I. An analysis of factors used in search engine ranking. Adversarial Information Retrieval on the Web, 48 (2005).
4. Brutlag, J. Speed matters for google web search, 2009 (http://www.isaacsunyer.com/wp-content/uploads/2009/09/test_velocidad_google.pdf, accessed Jul 31, 2013).
5. chandlernguyen.com, Search Engine Market Share Mar 2013, (http://www.chandlernguyen.com/2013/03/search-engine-market-share-mar-2013.html, accessed Jul 31, 2013).
6. Charlton, G. Eight second rule for e-commerce websites now halved, 2006 (http://econsultancy.com/uk/blog/500-eight-second-rule-for-e-commerce-websites-now-halved, accessed Jul 31, 2013).
7. comScore. comScore reports global search market growth of 46 percent in 2009, 2010 (http://www.comscore.com/Press_Events/Press_Releases/2010/1/Global_Search_Market_Grows_46_Percent_in_2009, accessed Jul 31, 2013).
8. Cutts, M. Google incorporating site speed in search rankings, 2010 (http://www.mattcutts.com/blog/site-speed/, accessed Jul 31, 2013).
9. Dover, D. and Dafforn, E. Search engine optimization (seo) secrets, 2011.
10. Enge, E., Spencer, S., Fishkin, R. and Stricchiola, J. C. The art of SEO. O'Reilly Media, Inc., 2010

11. Evans, M.P. Analysing google rankings through search engine optimization data. Internet research 17, 1 (2007), 21-37.
12. Frydenberg, M., and Miko J. Taking it to the top: A lesson in search engine optimization. Information Systems Education Journal 9, 1 (2011), 24.
13. Goldman, E. Search engine bias and the demise of search engine utopianism. Yale JL & Tech 8 (2005), 188-200.
14. Google. Webmasters tools, 2012 (http://www.google.com/webmasters/tools/, accessed Jul 31, 2013).
15. Google Code. Let's make the web faster, 2012 (http://code.google.com/speed/articles/, accessed Jul 31, 2013).
16. Google Webmaster Central Blog. Your site's performance in webmaster tools, 2009 (http://googlewebmastercentral.blogspot.com/2009/12/your-sites-performance-in-webmaster.html, accessed Jul 31, 2013).
17. Google Webmaster Central Blog. Using site speed in web search ranking, 2010 (http://googlewebmastercentral.blogspot.com/2010/04/using-site-speed-in-web-search-ranking.html, accessed Jul 31, 2013).
18. Google Webmaster Tools Help. Crawl stats, 2012 (http://support.google.com/webmasters/bin/answer.py?hl=en&answer=35253, accessed Jul 31, 2013).
19. Google Webmaster Tools Help. Google crawlers, 2012 (http://support.google.com/webmasters/bin/answer.py?hl=en&answer=1061943, accessed Jul 31, 2013).
20. Google Webmaster Tools Help. Search engine optimization (seo), 2012 (http://www.google.com/support/webmasters/bin/answer.py?answer=35291, accessed Jul 31, 2013).
21. Grappone, J., and Couzin, G. Search engine optimization: an hour a day. Sybex, 2011.
22. Bixby J.: 13 Questions (and Answers) About Google, Site Speed, and SEO, (http://www.seomoz.org/ugc/13-questions-and-answers-about-google-site-speed-and-seo, accessed Jul 31, 2013).
23. Gutzman, A. Living with the eight seconds rule, 2000 (http://www.ecommerce-guide.com/solutions/technology/article.php/518111, accessed Jul 31, 2013).
24. IAB. IAB internet advertising revenue report. 2012.
25. Jansen, B.J., Sobel, K. and Zhang, M. The Brand Effect of Key Phrases and Advertisements in Sponsored Search, International Journal of Electronic Commerce, 16, 1 (Fall 2011), 77-106.
26. Jones, T. Both sides of the digital battle for a high rank from a search engine. Association for Computing Machinery New Zealand Bulletin, 1, 2 (2005) 1-6.
27. Kohavi, R. and Longbotham, R. Online experiments: Lessons learned. Computer 40, 9 (2007) 103-105.
28. Kowalski, P. and Król, D. An approach to evaluate the impact of web traffic in web positioning. Agent and Multi-Agent Systems: Technologies and Applications, Lecture Notes in Computer Science, 6071 (2010) 380-389.
29. Marszałkowski, J.M. Analysis of the availability of network services. Master thesis, Poznań University of Technology. Unpublished (2010).
30. McDonald, J.H. Handbook of biological statistics. Sparky House Publishing, 2009.
31. McGee, M. It's official: Google now counts site speed as a ranking factor, 2010 (http://searchengineland.com/google-now-counts-site-speed-as-ranking-factor-39708, accessed Jul 31, 2013).

32. Rachitsky, L. Your sites performance now affects your google search ranking, 2010 (http://www.transparentuptime.com/2010/04/your-sites-performance-now-affects-your.html, accessed Jul 31, 2013).
33. Rodgers, J.L. and Nicewander, W.A. Thirteen ways to look at the correlation coefficient. American Statistician, 42, 1 (1988) 59-66.
34. Rutz, O.J. and Bucklin, R.E. A Model of Individual Keyword Performance in Paid Search Advertising, 2007 (http://ssrn.com/abstract=1024765, accessed Jul 31, 2013).
35. Seda, C. Search Engine Advertising: Buying Your Way to the Top to Increase Sales. New Riders Publishing, Thousand Oaks, CA, USA, 2004.
36. Sen, R. Optimal search engine marketing strategy. International Journal of Electronic Commerce, 10, 1 (2005) 9-25.
37. Sedigh, A.K. and Roudaki, M. Identification of the dynamics of the google's ranking algorithm. 13th IFAC Symposium On System Identification, Citeseer (2003).
38. SEOmoz. Search engine ranking factors, 2011 (http://www.seomoz.org/article/search-ranking-factors, accessed Jul 31, 2013).
39. Serrano-Cinca C., Fuertes-Callen Y., Gutierrez-Nieto B. Internet positioning and performance of e-tailers: An empirical analysis (2010) Electronic Commerce Research and Applications, 9 (3), pp. 237-248.
40. Skiera B., Eckert J., Hinz O. An analysis of the importance of the long tail in search engine marketing (2010) Electronic Commerce Research and Applications, 9 (6) , pp. 488-494.
41. wired.com. Ted 2011: The "panda" that hates farms, 2011 (http://www.wired.com/epicenter/2011/03/the-panda-that-hates-farms/all/1, accessed Jul 31, 2013).
42. Yang S. and Ghose, A. Analyzing the Relationship Between Organic and Sponsored Search Advertising: Positive, Negative, or Zero Interdependence?. Marketing Science, 29, 4 (2010) 602-623.
43. Yao, S. and Mela, C.F. A Dynamic Model of Sponsored Search Advertising. Marketing Science, 30, 3 (2011) 447-468.