Relevance Effectiveness of Acronyms in Information Retrieval

Adrienn Skrop

Department of Applied Informatics, University of Pannonia, Nagykanizsa Campus, Zrínyi utca18, Nagykanizsa, 8800, Hungary, skrop.adrienn@uni-pen.hu

Abstract: Many people use the Web to obtain information from public institutions, companies, organizations etc.. These users typically use a Web search engine and the acronym of the institution to get to the desired homepage. Users prefer using acronyms because they are easy to remember and are extensively used in media. This paper is concerned with studying the relevance effectiveness of Hungarian higher educational acronyms on the Web. The goal of this study is to evaluate and compare the ability of the acronyms of Hungarian institutions' names to find the home page of their own institutions when being used as queries in Web searching.

Keywords: Web retrieval, relevance effectiveness, acronyms

1 Introduction

The Web has become one of the most popular and important Internet applications both for users and information providers. One important category of information stored in Web pages is the category of higher educational institutions.These Web pages of institutions and organisations are of great interest to a large mass of users. The primary aim of a user wanting to obtain information from a specific institution is to get to the home page of that institution as easily and quickly as possible. The typical user scenario is as follows: (i) select a Web search engine, (ii) enter the acronym (or full name) of institution as a query, (iii) examine the first page of the hit list. Because acronyms of higher edicational institutions' names are commonly and very often used, e.g. during the admisson prucedure, the aim of this paper is to investigate the relevance effectiveness of these acronyms.

2 Problem statement

Information retrieval (IR) is a broad area of computer science. It deals with the representation, storage, organization of, and access to items, mainly documents. Its primary objective is to find unstructured documents from large collections that satisfies a user information need [1]. Nowadays IR is an everyday activity when people use a web search engine.

The evaluation of IR systems plays an essential role. The most important type of evaluation of IR systems is retrieval effectiveness evaluation. Traditional information retrieval evaluations and early Web experiments measures how well a given system or algorithm can match, retrieve and rank documents that are relevant to user's information need [2]. Of course, there are many studies on evaluating the effectiveness of Web search engines, for example [1][7], to quote just a few. However, not only this kind of task is a typical WWW search task. Navigational task is another activity. Its typical application is the home page finding task. The home page finding problem is one where the user wants to find a particular site and the query names the site. Home page finding queries typically specify entities such as people, companies, departments and products.

In [2], the home page finding problem is addressed from a user's viewpoint. Measures were elaborated to evaluate the home page identification capability of queries on the Web. The Pseudo Precision and Mean Reciprocal Rank measures were elaborated using the hazard rate function of the Mathematical Theory of Reliability and the Mean Reciprocal Rank (MRR) measure of traditional home page finding.

The most extensively used evaluation measure related to home page finding task is MRR. It gives an indication of how many low value results a user would have to skip before reaching the correct answer [5]. The MRR measure is commonly used when there is only one correct answer.

The Mathematical Theory of Reliability develops methods of evaluating the reliabilities of devices and introduces various quantitative indices for measures of devices performance [4]. The reliability of a device is defined to be the probability of performing its purpose adequately for the time intended. The most important measure of reliability is the failure rate function $\lambda(t)$. The empirical value of failure rate is the number of failures that can be expected to take place over a given unit of time.

The present paper is concerned with studying and comparing the usefulness of the acronyms of Hungarian higher educational institutions present on the Web. The working hypothesis is that higher educational acronyms are more effective than ten years ago.

The rest of the paper is organized as follows. Section 3 describes the experiment, including the test data. Section 4 gives an overview over the implemented

methodology and measures. Section 5 presents and discusses the results. Section 6 shows the conclusions.

3 Experimental Setting

The goal of the experiments was to evaluate and compare the relevance effectiveness of acronyms of Hungarian higher educational institutions on the Web. The first experiment [1] was carried out in 2004, whilest the second experiment was carried out in 2016.

3.1 Experiment 1

We identified 66 Hungarian higher educational institutions that have acronyms and are present with their own Web site on the Web [8][9][10]. Six Web search engines were used to evaluate the effectiveness of the acronyms. Table 1 presents the selected search engines.

Table 1
List of selected search engines: the first three are Hungarian search engines; the next three are general search engines

Name of the Search Engine	URL of the Search Engine
Heuréka	http://www.heureka.hu
AltaVizsla	http://www.altavizsla.hu
Ariadnet	http://www.ariednet.hu
Google	http://www.google.com
Metacrawler	http://www.metacrawler.com
AltaVista	http://www.altavista.com

In 2004, Heuréka, AltaVizsla and Ariadnet were the most frequently used Hungarian search engines (they were hosted and operated in Hungary) in Hungary which primarily indexed and searched Hungarian Web pages. They were preferred by most Hungarian users, who were lay people and have language difficulties when trying to use search engines in another language. However, three well-known general search engines (Google, Metacrawler, AltaVista) were also used because, on the one hand, they are preferred by the computing society, and, on the other hand, non-Hungarian speaking people might want to find out information on Hungarian institutions (for example, when they plan to travel to Hungary, or if they plan to study in Hungary, or if they live in Hungary).

3.2 Experiment 2

We identified 34 Hungarian higher educational institutions that have acronyms and are present with their own Web site on the Web [10]. Four Web search engines were used to evaluate the effectiveness of the acronyms. Table 2 presents the selected search engines.

Table 2
List of selected search engines in 2016

Name of the Search Engine	URL of the Search Engine
Google	https://www.google.com
Bing	https://www.bing.com/
Baidu	http://www.baidu.com/
Yahoo!	https://www.yahoo.com/

According to the Netmarketshare[1] statistics, the above search engines have the largest market share in 2016. The statistics are as follows:

Table 3
Search engine market share statistics

Search Engine	Total Market Share
Google	75,20%
Bing	8,40%
Baidu	7,69%
Yahoo!	6,88%

4 Methodology

Both of the experiments were implemented in the same way. Lists were compiled containing the full name, home page URL, and acronym for every institution. The experiments were carried out by entering each acronym to each of the selected search engines, and evaluating the first ten hits according to predefined categories and effectiveness measures.The examination of the first ten hits is suggested because users typically do not examine more links [3].

[1] https://netmarketshare.com/

Every link was assigned to exactly one of the following two categories:

Category 1: link to the home page of the institution. This is the Web page that is desired to be retrieved when the institution's acronym is used as query.

Category 2: irrelevant link. No link to the desired home page.

In the following, the measures that were used to evaluate the effectiveness of the acronyms are presented.

4.1 Pseudo-precision

For every acronym *a*, a measure called *pseudo–precision*, denoted by Π_a, is defined as follows:

$$\Pi_a = \frac{r_a}{N} \tag{1}$$

where *N* is the total number of search engines used, and r_a is the number of search engines that return category 1 links. Pseudo-precision is an analogue of the classical precision measure, and means the proportion of search engines for which the acronym proves useful.

4.2 Mean Reciprocal Rank

For every acronym *a* and every search engine *i*, a measure called *reciprocal rank*, denoted by RR_{ia}, is defined first as follows:

$$RR_{ia} = \begin{cases} \dfrac{1}{r_{ia}} & \text{category1\_link\_in\_ position\_ } r_{ia} \\ 0 & \text{no\_link\_in \_category1} \end{cases} \tag{2}$$

where r_{ia} is the rank of the link in the hit list, and κ is a penalty factor.
A *mean reciprocal rank*, denoted by MRR_a, for every acronym is defined as follows:

$$MRR_a = \frac{1}{N}\sum_{i=1}^{N} RR_{ia} \tag{3}$$

where *N* is the number of search engines used.

5 Results and Discussion

Figure 1 shows a pseudo–precision histogram of Experiment 1. These figures were obtained during our previous experiment detailed in [1] and are discussed here as a comparison. It can be seen that the majority, 67% of acronyms are not effective (pseudo-precision is less than 0,5); a few are very effective (pseudo-precision equals 1), and about 24% can be judged as being useful (pseudo-precision is more than 0.5).

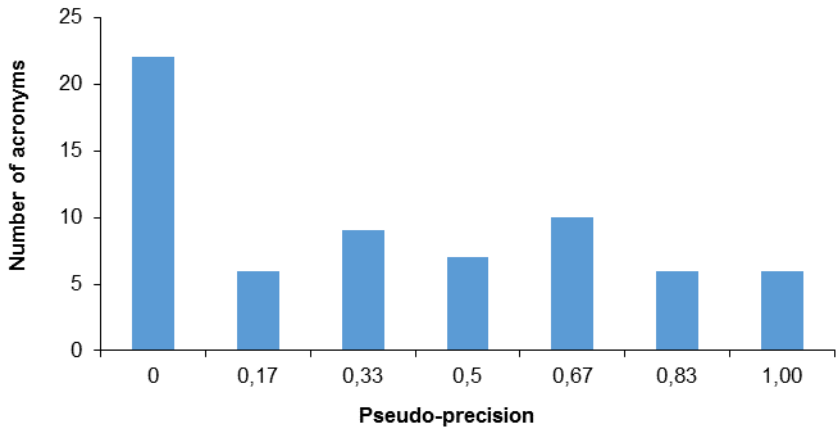

Figure 1
Pseudo –precision in Experiment 1

Figure 2 shows a pseudo–precision histogram of Experiment 2. These figures were obtained during this experiment.

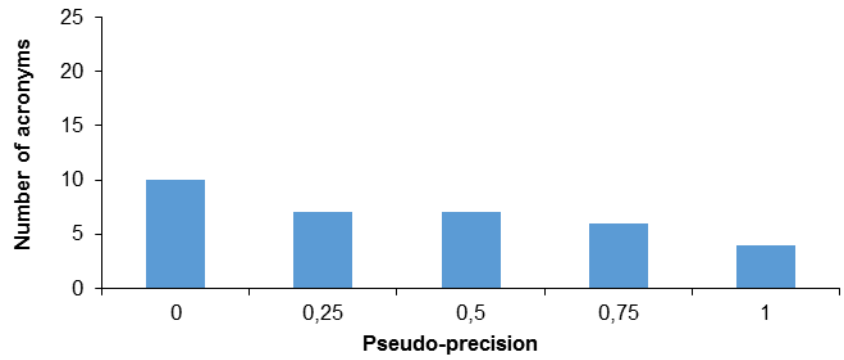

Figure 2
Pseudo –precision in Experiment 2

It can be seen that the majority of acronyms are not effective (pseudo-precision is less than 0,5); a few are very effective (pseudo-precision equals 1), and about 18% can be judged as being effective (pseudo-precision is more than 0.5).

Table 4 compares the distribution of acronyms over effectiveness categories. The usefulness categories are defined as follows:

- Not effective: pseudo–precision equals 0,
- Somewhat effective: 0 < pseudo–precision ≤ 0.5
- Effective: 0.5 < pseudo–precision < 1,
- Very effective: pseudo–precision equals 1.

The categories were defined so that both the most promising and also the worst cases can be compared. The categories do cover the whole interval of pseudo–precision.

Table 4

Percentage of acronyms over effectiveness categories

	Experiment 1	**Experiment 2**
Not effective	33%	29%
Somewhat effective	34%	42%
Effective	24%	17%
Very effective	9%	12%

The mean reciprocal rank histogram of acronyms of Experiment 1 are recalled from [1] and are presented in Figure 3 for comparison.

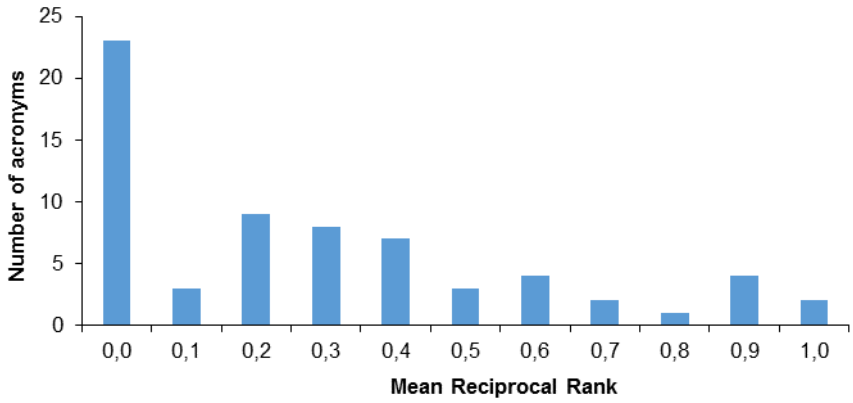

Figure 3

Mean Reciprocal Rank in Experiment 1

The mean reciprocal rank histograms of acronyms of Experiment 2 are obtained in this experiment.

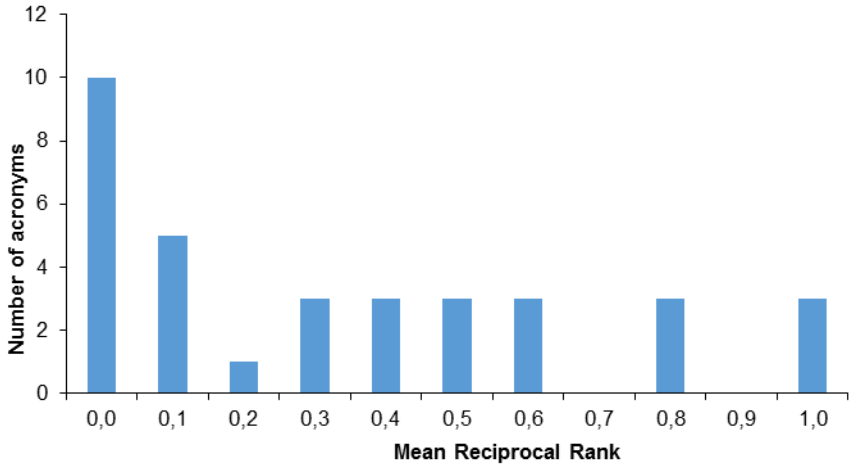

Figure 4
Mean Reciprocal Rank in Experiment 2

Table 5 shows the number and percentage of acronyms at different *MRR* intervals. The intervals represented in Table 5 are half-closed, and only the upper endpoints are included.

Table 5
The number and percentage of acronyms at different *MRR* intervals. The intervals represented are half-closed, and only the upper endpoints are included.

MRR	**Number of acronyms in Experiment 1**	**Percentage**	**Number of acronyms in Experiment 2**	**Percentage [%]**
0	23	35%	10	29%
0 – 0.1	3	5%	5	15%
0.1 – 0.2	9	14%	1	3%
0.2 – 0.3	8	12%	3	9%
0.3 – 0.4	7	11%	3	9%
0.4 – 0.5	3	5%	3	9%
0.5 – 0.6	4	6%	3	9%
0.6 – 0.7	2	3%	0	0%
0.7 – 0.8	1	2%	3	9%
0.8 – 0.9	4	6%	0	0%
0.9 –1	2	3%	3	9%

Average values of pseudo-precisions and mean reciprocal ranks are shown in Table 4.

Table 4
Average effectivenes measure values

	Experiment 1	**Experiment 2**
Average Pseudo precision	0,38	0,40
Average MRR	0,26	0,31

This poor performance of acronyms can be seen as a noteworthy situation. Analysing possible causes, the following two causes were identified. Poor Web page design of home pages and sites is one of the causes. Apart from content information (using the acronym as content or meta-data) also page–related factors (format, placement of the title tag, frequency of keywords etc.) and overall Web page design contribute to search engine ranking. The usage of title tags, fonts, character sizes, colours as well as of content need be considerably revised and improved by webmasters.

Conclusions

The effectiveness of the acronyms of Hungarian higher educational institutions to identify institutions was evaluated on the Web using search engines. The effectivness of acronyms from 2004 was used as a comparison. The results showed that the majority of the acronyms were not effective in identifying their institutions. The results refuted the working hypothesis that acronyms of higher educational institutions were more efective than more than ten yars ago. This means that they still fail to fulfil their roles of identifying their institutions, on the otther hand webmasters should seek ways to improve on the situation by a more careful design of home pages taking into account the different ways in which different search engines index and rank Web pages. These are relevant factors where acronyms of institutions' names are commonly and very often used in both media and by people in everyday life.

References

[1] A. Skrop and S. Dominich, (2004). Measuring the Identification Capability of Acronyms on the World Wide Web: a Comparative Study. *Journal of Web Engineering.* Vol. 3. No. 3 & 4, pp. 200-215. Rinton Press, ISSN 1540-9589.

[2] A. Skrop, (2006). New Methods in Web Information Retrieval Effectiveness. *PhD Theses.*

[3] A. Spink et al. (2002): From E-Sex to E-Commerce: Web Search Changes. *IEEE Computer* 35(3), pp. 107-109.

[4] B. V. Gnedenko et. al., (1969). Mathematical Methods of Reliability Theory, Academic Press, New York.

[5] C. Oppenheim, A. Morris, and C. McKnight, (2000). The evaluation of WWW search engines," *Journal of Documentation*, 56(2), pp. 190–211.

[6] D. M. Christopher, et. al., (2008). Introduction to Information Retrieval. Cambridge University Press, New York, NY, USA.

[7] H. V. Leighton and J. Srivastava, (1999). First 20 Precision among World Wide Web Search Services (Search Engines)," *Journal of the American Society for Information Science*, 50(10), pp. 870–881.

[8] http://egyetem.lap.hu/

[9] http://foiskola.lap.hu/

[10] http://www.felvi.hu/

[11] N. Craswell and D. Hawking, (2001). Overview of the TREC-2001 Web Track. In Vorhees, E. and Harman, D. K. (eds.). Proceedings of the 10th Text Retrieval Conference (TREC-2001), pp. 61-68. NIST special publication 500-250.

[12] S. Mizzaro, (1997). Relevance: The Whole History. *JASIS 48, 9*. pp. 810-832.