Are Semantically Related Links More Effective for Retrieval?

Marijn Koolen[1] and Jaap Kamps[1,2]

[1] Archives and Information Studies, University of Amsterdam,The Netherlands
[2] ISLA, University of Amsterdam,The Netherlands

Abstract. Why do links work? Link-based ranking algorithms are based on the often implicit assumption that linked documents are semantically related to each other, and that link information is therefore useful for retrieval. Although the benefits of link information are well researched, this underlying assumption on why link evidence works remains untested, and the main aim of this paper is to do exactly that. Specifically, we use Wikipedia because it has a dense link structure in combination with a large category structure, which allows for an independent measurement of the semantic relatedness of linked documents. Our main findings are that: 1) global, query-independent link evidence, is not affected by the semantic nature of the links, and 2) for local, query-dependent link evidence, the effectiveness of links *increases* as their semantic distance *decreases*. That is, we directly observe that links between semantically related pages are more effective for ad hoc retrieval than links between unrelated ones. These findings confirm and quantify the underlying assumption of existing link-based methods, which sheds further light on our understanding of the nature of link evidence. Such deeper understanding is instrumental for the development of novel link-based methods.

Keywords: Links, Semantic Relatedness, Effectiveness, Wikipedia

1 Introduction

Link-based ranking algorithms such as spreading activation [2], relevance propagation [19], HITS [8] and SALSA [12], use the assumption that linked documents have related content. For example, Picard and Savoy [16] say "the implicit reasoning made in spreading activation (SA) technique is the following: a link from a document d_1 to a document d_2 is *evidence* that their content is similar or related, such that if d_1 is relevant to a given request, d_2 may also be relevant."

So far, this assumption has remained implicit, because it is hard to measure the semantic relatedness of linked documents independent from the feedback of a retrieval system given a search query. Wikipedia allows us to explicitly measure the semantic relatedness of linked documents independently, and, with the INEX Wikipedia Ad Hoc test-collections since 2006, also allows us to study its impact on the effectiveness of link evidence for retrieval. Kamps and Koolen [7] found that Wikipedia links behave very much like links on the larger Web. Wikipedia being a part of the Web, we expect our findings to be generally applicable.

The algorithms mentioned so far are all query-dependent methods, but there are also query-independent methods, which might rely less on the semantic nature of links. In our analysis, we make a distinction between algorithms that use global, query-independent evidence, such as PageRank [15], and local, query-dependent algorithms such as HITS, which use the links between a subset of documents retrieved for a given topic in a query-dependent way. In this paper, we use the term local link evidence to refer to the links between the top 100 results for a given query.

Najork [14] showed that for large-scale Web retrieval, SALSA is more effective than PageRank, indicating that local link evidence is more effective for retrieval than global link evidence. A similar observation was made by Kamps and Koolen [7], who compared the effectiveness of global and local link degrees on Wikipedia ad hoc retrieval. These findings suggest that local link evidence reflects not only document importance, but also topical relevance [9]. Note that the query-dependent set of links is a proper subset of the global link structure. Evidently, some, but not all, links are useful for retrieval. This confronts us with the question:

– Are links between semantically related documents more effective for ad hoc retrieval than links between unrelated ones?

Wikipedia has a complex category structure, providing us with a hierarchical semantic classification of the articles. Thus, we can see whether a link connects two documents in the same category—in which case there is a clear semantic aspect to the link—or between two documents belonging to very different categories. We can also use category hierarchy to measure the semantic distance between two documents, which gives us a more fine-grained measure of semantic relatedness.

By filtering links based on their distance—removing the longest semantic distance links or the shortest semantic distance links—we can study the impact of the semantic nature of links on the effectiveness of link evidence. But filtering not only changes the semantic nature of link evidence but also the quantity. We therefore compare semantic filtering of links against a random filter. This leads to the more specific research questions:

– How is the link structure related to the categorical organisation in Wikipedia?
– How does semantic filtering of links affect the impact of link evidence on retrieval?

The rest of this paper is organised as follows. We first discuss related work in Section 2, and describe the category structure and look at the semantic relatedness of documents in Section 3. In Section 4 we address the issue of measuring semantic relatedness using the Wikipedia category structure. We then analyse how linked documents in the global and local link graph are distributed over the semantic relatedness measure in Section 5. Then, in Section 6 we describe experiments with filtering links using the category structure and finish with conclusions in Section 7.

2 Related Work

Link-based ranking algorithms like spreading activation [2], relevance propagation [19] and HITS [8] all use the implicit assumption that linked documents tend to be related to each other and therefore, that link information is potentially useful for retrieval. Consider the expansion step of the HITS algorithm: Starting with the highest ranked results for a query, this set is expanded by pages connected to those results to make sure the most important authorities on the search topic are included in the expanded set. The first part of this assumption was confirmed by [4], who showed that links on the Web tend to connect pages with topically related content.

The benefits of link information for information retrieval have been well-researched. A recent, large-scale evaluation of well-known algorithms such as PageRank [15], HITS and SALSA [12] was conducted by Najork [14]. On a large Web crawl and some 28,000 queries, he found that any link-based algorithm, including simple in-degree counts clearly outperform a random ordering of the same results. Link information is useful for ranking documents. However, these results do not explain why link information is useful.

Kurland and Lee [10, 11] showed that generating links based on document similarity can help improve ad hoc retrieval effectiveness. Assuming that document similarity reflects some kind of semantical relation between documents, this result shows that links between semantically related documents are effective for retrieval. However, is does not show they are effective *because* they connect semantically related documents.

Measuring the semantic relatedness (SR) of documents can be done in many different ways. A good overview of SR methods can be found in [1]. For our purposes, the work of Strube and Ponzetto [20] is relevant, because they used the Wikipedia category and link structures to measure word relatedness, and found that path-based measures using the category hierarchy perform well. The effectiveness of simple path-based measures on the Wikipedia category structure for SR is supported by Zesch and Gurevych [21].

3 Wikipedia Category Structure

We use the INEX 2006 Wikipedia collection [5], consisting of over 650,000 documents. The Wikipedia category structure is more or less hierarchical—categories are linked to each other via hypernym/hyponym relations but can have multiple parent categories—and allows us to determine how semantically related two documents are, even when they are not assigned to the same category, based on distances between categories.

In Wikipedia, anyone can edit the category structure, and there is no standard way to create such taxonomies of categories: one person could introduce several intermediate levels between two categories where another would introduce none or only a few. Some of the relations are even cyclic in the sense that two categories can subsume each other. However, we assume that distances at

Table 1. Link degree and category size statistics of the Wikipedia collections.

	Description	min	max	mean	median	stdev
Category	# articles	0	4,534	16.82	4	56.87
	# children	0	1,581	1.69	0	8.55
	# parents	0	55	1.69	2	1.17
	distance	1	23	7.29	7	1.58
Article	# categories	1	41	2.20	2	1.64

the extreme ends of the distribution—the shortest and longest distances—can respectively be interpreted as semantically related and unrelated.

Some statistics on the category structure are given in Table 1. The category structure of the INEX 2006 Wikipedia collection contains 86,024 distinct categories. The top category in the hierarchy is called CATEGORIES, and almost all categories are connected to this top category via sub-category relations. There are 75,601 categories that contain articles and 10,423 categories that contain no articles but have only sub-categories. The mean number of articles per category is 16.82, but the median is lower (4), showing that the distribution is skewed. The mean number of parent and child categories is 1.69, but the median numbers of parent and child categories are 2 and 0 respectively. Thus, most categories are leaves in the category structure, connected to at least 2 broader categories. All articles in the collection are assigned to at least one category, with a mean (median) of 2.2 (2).

4 Measuring Semantic Relatedness

Given that Wikipedia is a collection of interrelated topics, we can view the category structure as a taxonomy of concepts and use methods from computational linguistics to measure SR. The easiest way is to make a distinction between a pair of documents belonging to the same category and a pair of documents belonging to different categories, and say that the former pair is *semantically similar* whereas the latter pair is not. To give insight into how links in Wikipedia are related to the category structure, we adopt a path-based measure that simply counts the number of edges along the shortest path between two concept nodes [17, 18]. The rationale behind this is that "the shorter the path from one node to another, the more similar they are" [18] and "the relatedness of two words is equal to that of the most related pair of concepts they denote" [1].

We opt for the path-based measure using the category hierarchy because it is simple, has proven to be reasonably effective in semantic relatedness evaluations [20, 21], and, as we will see in the next sections, is sufficient for our purpose of studying the impact of SR on the effectiveness of link evidence for retrieval. The category distance between two documents d_a and d_b is the minimum of the category distances between the categories of d_a and d_b:

$$dist_{\mathrm{cat}}(d_a, d_b) = \min_{c_i \ni d_a, c_j \ni d_b} dist_{\mathrm{cat}}(c_i, c_j)$$

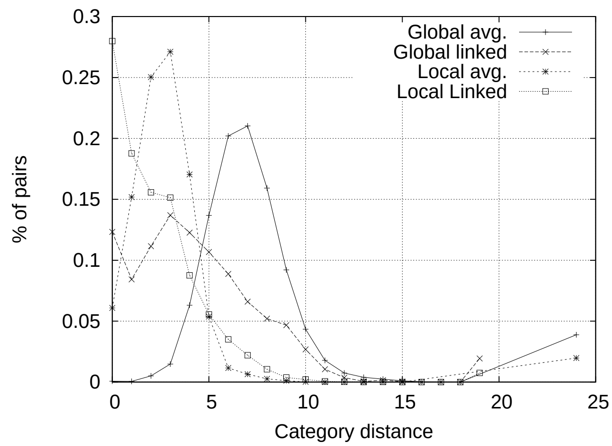

Fig. 1. Distribution of category distances between documents.

where $c_i \ni d_a$ are the categories to which d_a is assigned. The distance $dist_{\text{cat}}(c_i, c_j)$ between the two categories c_i and c_j is defined as:

$$dist_{\text{cat}}(c_i, c_j) = dist_{\text{cat}}(c_i, lso(c_i, c_j)) + dist_{\text{cat}}(c_j, lso(c_i, c_j)) \qquad (1)$$

where c_i and c_j are two categories, $lso(c_i, c_j)$ is the lowest super-ordinate (the lowest super category) of c_i and c_j and $dist_{\text{cat}}(c_i, lso(c_i, c_j))$ is the number of steps up the hierarchy from category c_i to $lso(c_i, c_j)$.

When we consider only categories connected to the top category CATEGORIES, the average shortest distance between two categories is 7.29 (median 7) and the maximum is 23. What is the average category distance between two pages? We randomly sampled one million pairs of documents and computed the shortest category distance between them. The distribution of category distances is shown in Figure 1 (the solid line, Global average). The distribution of the global pairs is roughly normally distributed, with a peak at distance 7, with 21% of the documents pairs. The bulk of the document pairs are at a category distance of 4–10, and very few document pairs are semantically close to each other. The right-most data points represent document pairs belonging to unconnected parts of the category structure. Among the pairs that are connected via the category structure the average distance is 6.61, which is slightly below the average distance between two categories, which is 7.29. We also computed the category distance between all document pairs in the local top 100 documents for the 221 topics. This resulted in 1,093,950 document pairs. The distribution has roughly the same shape but is shifted towards the smaller distances. In the top 100 retrieved documents, 6% of the document pairs share at least 1 category (distance 0), the most frequent distance is 3 and almost all pairs have a distance less than 6. Among the pairs that are connected via the category structure, the average distance is 2.56. The documents in the top 100 results are more semantically related to each other than in the overall collection.

5 Links and Categories

Now that we have chosen a method to measure semantic relatedness, we look at how the link structure is related to semantic relatedness. Again, one of the main assumptions underlying algorithms like HITS and relevance propagation [19] is that links are a signal that two documents are topically related to each other. But perhaps not *all* linked documents are topically related to each other. The Wikipedia category structure provides a manually created semantic organisation of the Wikipedia articles, with which we can quantify how related two articles are. How is the link structure related to the categorical organisation in Wikipedia? We look at the shortest category distance between linked articles. The distribution of links over shortest category distance is given in Figure 1 and is shown both globally and locally over the top 100 retrieved results. The local top 100 results are based on the 221 Ad Hoc topics and associated relevance judgements of the INEX Ad Hoc test-collections of 2006–2007 [6, 13]. The baseline retrieval system is described in the next section.

In the global link structure, around 12% of the links connect two articles sharing at least one category—from here on referred to as *within-category links*, as opposed to *cross-category links*, which connect documents that share not a single category. The most frequent distance is 3 steps, above which the frequency gradually drops to almost 0 at 12 steps. There is a small peak again at the end, for the links between articles assigned to unconnected categories.

Linked documents tend to be more semantically related to each other than randomly paired documents and share a category much more often. The median category distance of the linked documents is 4 while the median of the randomly paired documents is 7. Among the linked documents that are connected via the category structure, the average distance is 4.04, compared to 6.60 for the randomly sampled pairs. *There is a clear relation between global links and semantic relatedness.* However, compared to the documents in the top 100, the linked documents share a category more often but are also more frequently separated by greater semantic distances. Within the top retrieved results, the global link evidence has a weaker semantic signal than the text evidence.

The category distance distribution over the local links is based on 63,435 links between the documents in the top 100 results of the 221 topics (5.8% of all possible pairs in the local sets). The local links show a very different distribution. Here, the 0 distance links are the most frequent and make up more than 25% of the link set, and the frequency drops monotonously over category distance, with almost no pairs beyond 8 steps. There is a small set of links between articles assigned to unconnected categories. This means there is a clear relation between local link evidence and SR. In the query-dependent link set we more frequently find links between articles that are semantically similar. This is not surprising, because each article appears in the local set because it shows similarity with the search query and therefore also with the other documents in the local set. However, the average distance of the linked document pairs is 2.22 while over the entire local set the average is 2.56. In the top 100 results of a given query, the local links provide a stronger signal that two documents are semantically related than the text evidence.

How is the link structure related to the category structure? There is a clear relation between global links and SR. However, this semantic signal is weaker than the text evidence in the top retrieved documents. In the local set, pages that are linked tend to be more semantically related than pages that are not linked. Is the semantic nature of links also related to their effectiveness for information retrieval? This question is addressed in the next section.

6 Semantic Relatedness and Effectiveness of Links

By zooming in on the top ranked retrieval results, we filter the link graph on the search topic and end up with links between semantically related pages. The global link graph contains the same links but also many more links between semantically unrelated pages. How is the impact of link evidence related to the semantic nature of links? We use the category structure to filter links and thereby control the semantic nature of link evidence. What happens to the impact of link evidence if we remove the within-category links? Does link evidence become less effective? What happens when we remove only the longest distance links?

Our baseline run is a standard language model run with linear smoothing ($\lambda = 0.15$) and a document length prior $P_{length}(d) = |d|/\sum_{d' \in D} |d'|$, where d and d' are documents in collection D. The length prior promotes longer documents and improves MAP from 0.2561 to 0.3157.

To study the effectiveness of link evidence, we look at link degrees, which have proven to be very competitive compared to more complex algorithms like PageRank and HITS [7, 14], and are simpler to compute. As in [7], we concentrate on the top 100 results for each topic. We experimented with in-degrees, out-degrees and their union (treating links as undirected), and found that for global link evidence, the in-degree is more effective than out-degree or their union. For local link evidence, in- and out-degree are equally effective, but their union is more effective. As global, query-independent evidence, links are more effective in one direction, which suggests they provide evidence of document importance. As local, query-dependent evidence, links are effective in both directions, suggesting their evidence is symmetric, and might reflect semantic relatedness, which is symmetric as well. For lack of space, we restrict our discussion to the global in-degrees and the local union degrees.

To show how the semantic nature of links affects their impact on effectiveness, we use two filtering methods: one where we remove the shortest semantic distance links (the SD filter), effectively degrading the semantic nature of the link graph, and one where we remove the longest semantic distance links (LD filter), effectively improving the semantic nature of the link graph. We filter links based on the path length distance measure described above. In the first filtering step the SD filter removes the links at distance 0, in the second step the links at distance 1, etc. The LD filter first removes the links between pages unconnected

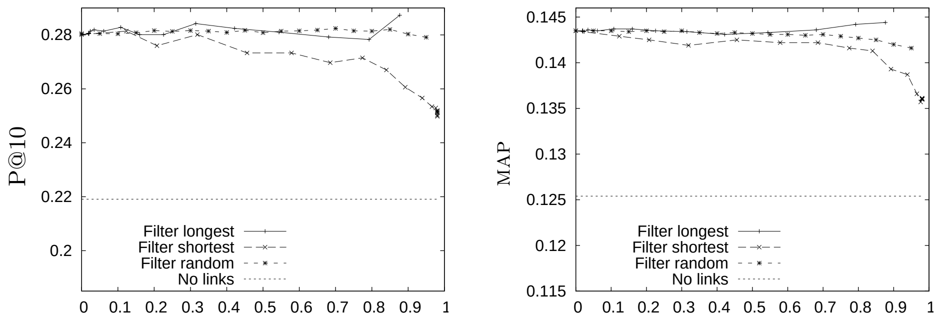

Fig. 2. The impact of filtering links on the effectiveness of ranking the top 100 results by global in-degree. The x-axis shows the percentage of links removed.

to each other via the category structure. In the second step the LD filter removes links at the largest distance (18 steps, see Figure 1), etc.

Note that by filtering we not only affect the semantic nature of the link graph, but also the link quantity. For comparison, we also look at the impact of randomly filtering links. We do this by assigning a random value between 0 and 1 to each page in the collection and sampling $n\%$ of the pages by selecting all pages with a value below $\frac{n}{100}$. The degree distribution of an $n\%$ sample is determined by the random assignment of the values, so repeating the experiment can result in different distributions. Therefore, the values reported are the averages over 20 iterations. If we randomly remove links from the graph, we would expect that the degrees change uniformly. That is, all pages are affected in the same way.

We look at the top 100 results retrieved by the text retrieval baseline and compare the ranking based on link evidence against a random ordering of documents. This shows whether link evidence is has any potential value. The impact of filtering on the effectiveness of the global in-degrees is shown in Figure 2. The x-axis shows the percentage of links filtered.

The left figure shows the impact on P@10. The Random and LD filters have little impact on the in-degrees, but removing the shortest distance links hurts performance. Performance stays well above that of random ordering though. On MAP (right figure) the impact of filtering is similar to the impact on P@10. The in-degree performance slightly improves with only the within-category links and drops with only the 10% longest distance links. From these observations we learn that filtering has little impact on the global degrees, probably because the link graph is very rich and the high-degree pages are very robust against filtering.

Although filtering does not improve performance for the in-degrees, we note that using global out-degrees (not shown) is not effective—no better than random—unless we filter out the longest distance links. A large number of links to semantically related pages signals that a page is a good hub for a particular topic.

The impact of filtering on the local degrees is shown in Figure 3. Note that without filtering, local links are far more effective than global links. Here, random filtering has a bigger impact. The local link graph is already filtered on the search topic and has far fewer links. Further filtering flattens the degree distribution

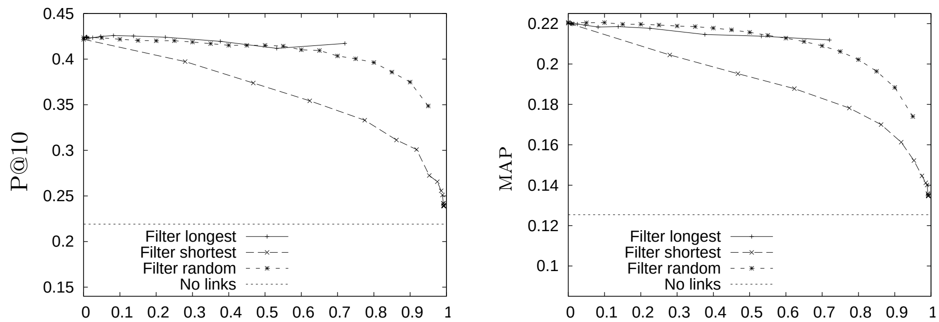

Fig. 3. The impact of filtering links on the effectiveness of ranking on local union degrees. The x-axis shows the percentage of links removed.

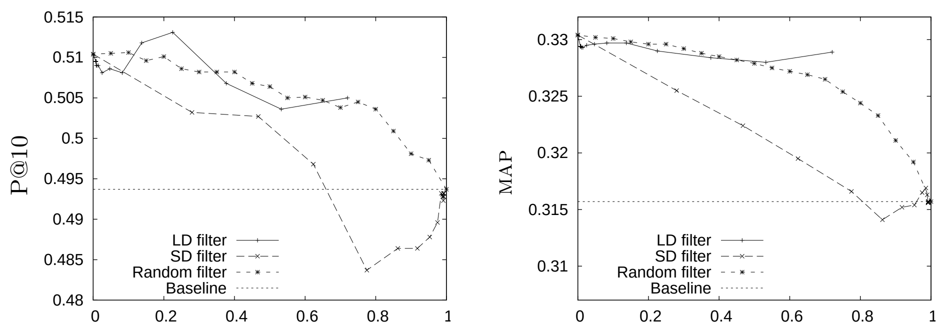

Fig. 4. The impact of filtering links on the effectiveness of ranking on local union degree and text evidence. The x-axis shows the percentage of links removed.

even more. If we remove the shortest distance links first, performance drops faster than with random filtering, while if we remove the longest distance links first, performance remains stable. The shortest semantic distance links are the more effective links. If we want to improve ad hoc search by exploiting link evidence, we need links between semantically related pages. Another important thing to note is that filtering on the category structure does not make local link evidence more effective. Zooming in on the highest ranked retrieval results already gets rid of most links between unrelated pages. Further filtering is not needed.

What happens to the performance of link evidence in combination with the content-based score when we filter links? There are many ways to combine content and link evidence (see, e.g. [3]). We experimented with several combination methods, such as using the log of the degrees or prior probabilities trained on the relevance data, instead of the degrees themselves. Although the impact of filtering is similar for the different methods, we found that the most effective combination is to multiply the document score from the baseline model by the local union degree plus 1 (so that documents with no links keep their original document score). That is, the final score is $S(d) = S_{base}(d) \cdot (1 + \text{degree})$, where $S_{base}(d)$ is the baseline score. The results are given in Figure 4. The baseline scores are the straight dotted lines. The local union degrees improve upon the baseline performance. With random filtering, both P@10 and MAP gradually drop as we remove more links. If we remove the SD links first, the improvement drops faster and the score even falls below that of the baseline. With the LD filter, the P@10 score fluctuates somewhat between 0.505 and 0.513, while the MAP score remains stable. With just the local within-category links, the improvement is the same as with all local links. Again, the links between the most semantically related documents are the most effective.

Note that filtering does not improve the effectiveness of local link evidence, which might be explained by the fact that the local link graph is already filtered on the search topic, which is a semantic filter in itself.

To summarise, global link evidence is very robust against filtering and its effectiveness seems unrelated to semantic relatedness. Local link evidence is more sensitive to filtering, partly because the graph is more sparse as is it already filtered on the search topic. But its effectiveness is directly related to the semantically relatedness of the linked documents.

7 Conclusions

This paper investigated the semantic nature of links, trying to answer whether links between semantically related pages are more effective for retrieval than links between unrelated ones. Our first research question was:

– How is the link structure related to the categorical organisation in Wikipedia?

Compared to a random sample of document pairs, linked documents tend to be more semantically related to each other and more often share a category, showing a clear relation between global links and semantic relatedness. However, within the top retrieved documents for a given query, the semantic signal of global link evidence is weaker than that of the textual evidence, providing an explanation why global link evidence is almost ineffective for topic relevance tasks. In the local set, pages that are linked tend to be more semantically related than pages that are not linked. Local link evidence is more clearly related to semantic relatedness and, even in the more topically focused set of top retrieved pages, links are a stronger signal that two pages are semantically related. This shows a difference in the semantic nature of global and local links. The semantic nature of link evidence changes as we zoom in on a subset of pages retrieved for a given query. Our second research question was:

– How does semantic filtering of links affect the impact of link evidence on retrieval?

Global incoming link evidence is robust against filtering links randomly or based on semantic distance, and only becomes less effective when the longest semantic distance links are left. The effectiveness of global link evidence is not determined by the semantic relatedness of linked documents. Local link evidence is less robust against filtering, becoming less effective when we remove links. Effectiveness drops as the number of short distance links drops. The effectiveness of local link evidence is thus, at least partly, determined by the semantic relatedness of linked documents. The step from a global link graph to a local link graph works as a semantic link filter. Many of the links between semantically unrelated pages are removed. This is an essential step in making link evidence useful for ad hoc search. Our hypothesis that link evidence for topical relevance is symmetric hinges on the semantic relatedness of linked pages.

Finally, our main aim was to investigate:

- Are links between semantically related documents more effective for ad hoc retrieval than links between unrelated ones?

When the aim of link evidence is to identify important documents, links between semantically related documents are not more effective than links between unrelated ones. When we make link evidence sensitive to the context of the search topic, the role of link evidence shifts to identifying topically relevant documents, and here links between semantically related documents are indeed more effective than links between unrelated ones.

More generally, our findings confirm the assumption that (query-dependent) link information is effective for retrieval because it signals the semantic relatedness of linked documents. This adds to our understanding of why link evidence works, which can help in developing better link-based ranking methods.

We did this analysis in Wikipedia because its category structure allows an independent measurement of the semantic relatedness of linked documents. The fact that the impact of link evidence in our experiments is similar to the results of other studies (e.g. local versus global evidence, [14]), and that Wikipedia links behave like general Web links [7], offers support that these findings generalise to the larger Web and hyperlinks in general.

In future work, we will extend our analysis to a general Web corpus and experiment with generating links based on content similarity (as done by Kurland and Lee [11]). We will also investigate better ways of combining link and text evidence, and look at the impact of weighting instead of filtering links.

Acknowledgments This work was generously supported by the Netherlands Organization for Scientific Research (NWO, grants # 612.066.513, 639.072.601, and 640.001.501).

References

[1] A. Budanitsky and G. Hirst. Evaluating wordnet-based measures of lexical semantic relatedness. *Computational Linguistics*, 32(1):13–47, 2006.
[2] P. R. Cohen and R. Kjeldsen. Information retrieval by constrained spreading activation in semantic networks. *Inf. Process. Manage.*, 23(4):255–268, 1987.
[3] N. Craswell, S. Robertson, H. Zaragoza, and M. Taylor. Relevance weighting for query independent evidence. In *Proceedings of the 28th annual international*

ACM SIGIR conference on Research and development in information retrieval, pages 416–423, New York, NY, USA, 2005. ACM. ISBN 1-59593-034-5.

[4] B. D. Davison. Topical locality in the web. In *Research and Development in Information Retrieval (SIGIR)*, pages 272–279, 2000.

[5] L. Denoyer and P. Gallinari. The Wikipedia XML Corpus. *SIGIR Forum*, 40(1): 64–69, June 2006.

[6] N. Fuhr, J. Kamps, M. Lalmas, S. Malik, and A. Trotman. Overview of the INEX 2007 ad hoc track. In *Focused access to XML documents: 6th International Workshop of the Initiative for the Evaluation of XML Retrieval (INEX 2007)*, volume 4862 of *Lecture Notes in Computer Science*, pages 1–23. Springer Verlag, Heidelberg, 2008.

[7] J. Kamps and M. Koolen. Is Wikipedia link structure different? In *Proceedings of the Second ACM International Conference on Web Search and Data Mining (WSDM 2009)*, pages 232–241. ACM Press, New York NY, USA, 2009.

[8] J. M. Kleinberg. Authoritative sources in a hyperlinked environment. *Journal of the ACM*, 46(5):604–632, 1999.

[9] M. Koolen and J. Kamps. What's in a link? from document importance to topical relevance. In *Proceedings of the 2nd International Conferences on the Theory of Information Retrieval (ICTIR 2009)*, volume 5766 of *LNCS*, pages 313–321. Springer Verlag, Berlin, Heidelberg, 2009.

[10] O. Kurland and L. Lee. Pagerank without hyperlinks: structural re-ranking using links induced by language models. In *SIGIR*, pages 306–313. ACM, 2005.

[11] O. Kurland and L. Lee. Respect my authority!: Hits without hyperlinks, utilizing cluster-based language models. In *SIGIR*, pages 83–90. ACM, 2006.

[12] R. Lempel and S. Moran. Salsa: the stochastic approach for link-structure analysis. *ACM Trans. Inf. Syst.*, 19(2):131–160, 2001.

[13] S. Malik, A. Trotman, M. Lalmas, and N. Fuhr. Overview of INEX 2006. In INEX, volume 4518 of *Lecture Notes in Computer Science*, pages 1–11. Springer, 2006.

[14] M. Najork. Comparing the effectiveness of hits and salsa. In *CIKM*, pages 157–164. ACM, 2007.

[15] L. Page, S. Brin, R. Motwani, and T. Winograd. The pagerank citation ranking: Bringing order to the web. Technical report, Stanford Digital Library Technologies Project, 1998.

[16] J. Picard and J. Savoy. Enhancing retrieval with hyperlinks: A general model based on propositional argumentation systems. *JASIST*, 54(4):347–355, 2003.

[17] R. Rada, H. Mili, E. Bicknell, and M. Blettner. Development and application of a metric on semantic nets. *IEEE Transactions on Systems, Man, and Cybernetics*, 19(1):17–30, 1989.

[18] P. Resnik. Using information content to evaluate semantic similarity in a taxanomy. In *Proceedings of the 14th International Joint Conference on Artificial Intelligence (IJCAI-95)*, pages 448–453, 1995.

[19] A. Shakery and C. Zhai. A probabilistic relevance propagation model for hypertext retrieval. In *CIKM*, pages 550–558. ACM, 2006.

[20] M. Strube and S. P. Ponzetto. Wikirelate! computing semantic relatedness using wikipedia. In *Proceedings of the Twenty-First National Conference on Artificial Intelligence*, July 2006.

[21] T. Zesch and I. Gurevych. Analysis of the wikipedia category graph for nlp applications. In *Proceedings of the TextGraphs-2 Workshop (NAACL-HLT 2007)*, pages 1–8, Apr 2007.