Information Retrieval by Inferring Implicit Queries from Eye Movements

David R. Hardoon, John Shawe-Taylor
Computer Science Dept.
University College London
Gower Street, London
UK, WC1E 6BT

Antti Ajanki, Kai Puolamäki, Samuel Kaski
Helsinki Institute for Information Technology
Laboratory of Computer and Information Science
Helsinki University of Technology
P.O. Box 5400, FI-02015 TKK, Finland

Abstract

We introduce a new search strategy, in which the information retrieval (IR) query is inferred from eye movements measured when the user is reading text during an IR task. In training phase, we know the users' interest, that is, the relevance of training documents. We learn a predictor that produces a "query" given the eye movements; the target of learning is an "optimal" query that is computed based on the known relevance of the training documents. Assuming the predictor is universal with respect to the users' interests, it can also be applied to infer the implicit query when we have no prior knowledge of the users' interests. The result of an empirical study is that it is possible to learn the implicit query from a small set of read documents, such that relevance predictions for a large set of unseen documents are ranked significantly better than by random guessing.

1 INTRODUCTION

Current information retrieval (IR) systems rely mostly on explicit, typed queries, combined with explicit feedback telling the system which of the search results were relevant. The relevance feedback is used to refine the query, and the search converges iteratively towards more relevant documents. The standard web search engines are simplified versions of this scheme; they take advantage of the large scale which allows inferring general interest of documents from link data.

A main disadvantage of the traditional IR paradigm is that formulating good textual queries is a challenging task, even for experienced users (Turpin & Scholer, 2006). The task is made even more difficult by the fact that the true interests of the users are often ambiguous even for the users themselves, and therefore it may be difficult to formulate the query explicitly. It would be ideal if the system could infer the interests of the users while they work, and then have some suggestions readily available when the users ask for help. We call this task *proactive information retrieval*.

A proactive information retrieval system would additionally solve the problem that giving explicit feedback is laborious. Such a system would use *implicit feedback* to infer relevance. Click-stream data is one form of implicit feedback. While useful and often readily available, it offers limited information about a user's interests and intentions (Joachims, Granka, Pan, Hembrooke, & Gay, 2005; Kelly & Teevan, 2003).

In this work, we propose to use implicit feedback from the observation of eye movements as an alternative source to infer the users' intentions. We combine the eye movements with the textual content of the documents in a novel way: *we use the eye movements to formulate an IR query*, which is then used to rank unseen documents with respect to their relevance to the current interests of the user.

While the traditional IR task is to find relevant documents given a query typed in by the user, we address the following tasks: (i) Construct a query from eye movements alone. (ii) Construct a query by combining information from implicit relevance feedback from eye movements *and* explicit relevance feedback.

We devise a controlled experimental setting, in which test subjects read through short text snippets searching for documents related to a given topic. During the reading the users' eye movements are recorded with an eye tracking system. We extract term-specific eye movement features for each document the user reads which are then used to predict importance of the term for the search task. The set-up is an extended version of earlier work (Puolamäki, Salojärvi, Savia, Simola, & Kaski, 2005).

We assume that there is a link between relevance or interest and eye movements, and that this link can be learned by observing the users' behaviour in search tasks where the ground truth (the true interest of the user) is known. Our main assumption is that this link between interest and eye movements is, to a reasonable extent at least, independent of the actual topic. Hence, we propose to construct an IR query for a previously unseen topic solely by observing the eye movements of the user during an IR task of finding documents on that topic.

This work is a feasibility study on whether there is any truth in this assumption, and whether it would be realistic to use eye movements to formulate queries in information retrieval. We assume that no explicit relevance feedback is available.

As another case study, we investigate whether combining the eye movements with document features would help in the standard IR task where explicit relevance feedback is available for a subset of the documents.

More formally, we work with a bag-of-words (BOW) representation of the documents. We use *term-specific eye movement features*, denoted collectively by $\mathbf{e}_t$ for term t, to predict the parameters of our query, denoted collectively by $\mathbf{w}$, as $w_t = f_\lambda(\mathbf{e}_t, \mathbf{s}_t)$, where $\mathbf{s}_t$ are possible query-independent parameters associated with the term t (e.g., document frequency of term t). The $f_\lambda()$ could in principle be any predictor with *term-independent* parameters specified by λ. We then use a query function $g_\mathbf{w}(d)$, where d is a BOW representation of a document, to rank unseen documents with respect to their predicted interest to the user.

The parameters λ of the predictor $f_\lambda()$ are learned in the training phase, where we know the ground truth (true interest of the user). Note that following our assumptions laid out above, we assume that the functional form of the predictor $f_\lambda()$ is independent of the actual interest of the user. Hence, we can use it to formulate query parameters $\mathbf{w}$ for previously unseen topics.

In the following we choose the query function $g_\mathbf{w}()$ to be a Support Vector Machine (SVM) based model, with term-specific parameters w_t, as we believe them to be most suitable for the proposed task. As a predictor $f_\lambda()$, which gives the parameters of the SVM model, we use some standard linear and non-linear regressors. As this is the first study of this kind, and we aim for robustness, we purposely use standard state-of-the-art machine learning methodologies. Developing methods tailored for the task is left for future work.

The paper is laid out as follows; In Section 2 we discuss the usage of eye movements in information retrieval as well as previous work. Section 3 provides a description of the data and how it was acquired. We discuss our proposed models in Section 4. In Section 5 we elaborate on the experiment set up while in Section 6 the results are given. Our final remarks and discussion are given in Section 7.

2 EYE MOVEMENTS IN INFORMATION RETRIEVAL

Use of eye movements in IR is a relatively new approach. Maglio, Barrett, Campbell, and Selker (2000); Maglio and Campbell (2003) introduced a prototype attentive agent application which monitors eye movements while the user views web pages, in order to determine whether the user is reading or just browsing. If reading is detected, more information of the topic is sought and displayed. The feasibility of the application was not however experimentally verified.

The eye movements were first used in an information retrieval task in (Salojärvi, Kojo, Simola, & Kaski, 2003; Salojärvi, Puolamäki, & Kaski, 2005). Discriminative hidden Markov models were applied to estimate the relevance of lines of read text, and the performance of the method was verified in a controlled experiment. A competition was subsequently set up, where the participants competed in predicting relevance based on the eye movements (Puolamäki & Kaski, 2006).

A prototype information retrieval system was introduced in (Puolamäki et al., 2005). The system used relevance information combined with collaborative filtering to seed out relevant scientific articles. This earlier prototype did not use the textual content of the documents at all.

3 DATA DESCRIPTION

The data set consists of about 500 Wikipedia documents from 25 different categories. The category titles were used as the queries. The training corpus is as depicted in Table 1. The Wikipedia documents were truncated to fit to the screen (11 lines, about a dozen words on each). We made sure that the content of each truncated document was sufficient for inferring its topic, by manually inspecting all documents. For testing the accuracy we had another data set of 244 non-truncated documents, 10 documents for each category, except for *Natural disasters* which had only 4 documents. The documents in the testing corpus were not used in any way during the training of the model. For the BOW representation, we define the dictionary to be the set of stemmed words occurring in the training corpus. Some frequent 'stop' words like 'of' and 'the' are omitted from the dictionary.

In the experiments the users were shown ten docu-

Table 1: Summary of the Training Corpus.

Search Topic	Number of Documents
Astronomy	23
Ball games	23
Cities	13
Court systems	23
Dinosaurs	17
Education	22
Elections	22
Family	18
Film	21
Government	21
Internet	23
Languages	22
Literature	23
Music	16
Natural disasters	21
Olympics	22
Optical devices	23
Postal system	23
Printing	23
Sculpture	20
Space exploration	23
Speeches	23
Television	23
Transportation	23
Writing systems	17

ments and they were asked, after they had viewed each document, to identify whether the document was relevant to a search topic they had been given beforehand (Fig. 1). Once the users had assessed the relevance of the document, they pressed any button, after which they reported the relevance. The next document showed up immediately. On average, half of the documents shown during each topic search were relevant. Each user repeated the task ten times, with different search topic and documents in each round. The user group consisted of ten post-graduate and senior researchers.

While the users were reading, their eye gaze position on the screen was recorded with a Tobii 1750 eye tracker. Tobii measures eye positions 50 times per second by illuminating both eyes with infrared LED and measuring the reflected light. The system is fairly robust to head movements. The eye tracker was calibrated in the beginning of the experiment for each user.

The gaze direction is an indicator of the focus of attention, since accurate viewing is possible only in the central fovea area (1–2 degrees of visual angle). The correspondence is not one-to-one, however, since the attention can be shifted without moving the eyes. The eye movement trajectory is traditionally divided into fixations, during which the eye is fairly motionless, and saccades, rapid eye movements from one fixation to another.

We extracted 22 eye movement features from the recorded eye movement data for each fixation, and 4 text features for the target words of the fixations. The eye movement features included, among others, the number of fixations on the word, relative and absolute fixation durations, and lengths of the saccades before and after the fixation. They are described in more detail in (Salojärvi, Puolamäki, Simola, et al., 2005). The text features were the number of characters in the word, relative position of the word in the document and on the line, and the inverse document frequency of the word. Fixation locations were found using the default values of parameters recommended for text-only stimuli by Tobii: if the recorded gaze points stayed inside a 20 pixel area (about 0.5 degrees of visual angle) for at least 40 ms, they were considered as one fixation. Every fixation was mapped to the closest word, unless the fixation occurred well outside any text, in which case it was discarded.

4 MODELS USED

Our main task is to formulate an IR query, using the eye movements as the only feedback signal. The query need not however be understandable by humans; in fact, it suffices to formulate the query in such a way that it can be used by a relevance predictor to predict relevance for new documents.

Our query is a *parameter vector* for a relevance predictor which uses the eye movements as inputs. For the training data, we need to know the correct solution; for this purpose we use the parameter vector, referred to as the *ideal weights*, of a predictor trained to classify according to known relevance labels which are available for the learning data. We need to estimate the ideal parameter vector and to construct a *regressor* that tries to predict the ideal weights given the eye movements.

Furthermore, we study with a straightforward *combination of the text and eye movements* whether the eye movements help in the relevance prediction task when explicit relevance feedback is also available.

4.1 SUPPORT VECTOR MACHINES

We use support vector machines (SVM) for two tasks; to compute the *ideal weights* and to predict relevance of unseen documents, and to combine eye movement and textual features in the IR task.

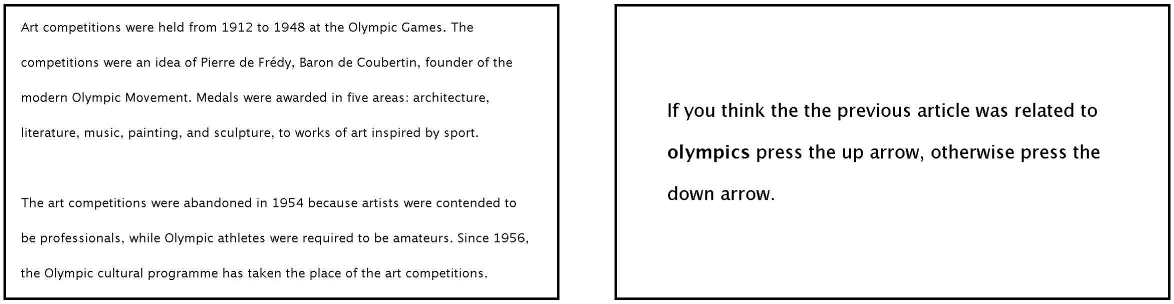

Figure 1: A sample document (on the left; it consists of two paragraphs) and the screen which the user gets after pressing a button (on the right).

4.1.1 Ideal Weights

We use an SVM[1] for each search topic, to predict the relevance labels (two classes: relevant or not) given by the user. The input is the BOW representation of the document. The SVM has one weight parameter for each term, and these weights are used as the ground truth or *ideal weights*.

The ideal weights are computed using one SVM per search topic. This procedure produces a weight value for each term in the dictionary; the weight represents the term's "fit" to the given search topic. We compute the SVM weights for all the non-truncated documents but restrict the dictionary to that of the truncated documents as there will never be feature vectors on terms that do not appear in the truncated documents. These SVM weights are the base in our work and we consider them to be the optimal baseline performance achievable.

We use a similar linear model, with weights given by the regressor described later in subsection 4.2, to predict the relevance of the unseen documents.

4.1.2 Combining Eye Movements and Explicit Feedback

We are interested in observing whether using eye movement and text features in conjunction with explicit document relevance information and document features would improve the classification accuracy in comparison with only using the document features. This task differs from the one above in that now explicit relevance feedback is assumed available.

Here we consider the measured eye movement information on the documents to be a second representation of the document.

When two representations of the same phenomenon are available kernel Canonical Correlation Analysis (KCCA) (Hardoon, Szedmak, & Shawe-taylor, 2004) has been shown to be an effective pre-processing step that can improve the performance of classification algorithms such as the SVM. The SVM-2K (Farquhar, Hardoon, Meng, Shawe-Taylor, & Szedmak, 2006) is a method that combines these two stages of learning (KCCA followed by SVM) into a single optimisation.

[1] All SVM's mentioned use the default setting of $C = 1$ and a linear kernel.

4.2 REGRESSION

We assume that there is a link between eye movements and the relevance of words within a document. Our proposed model aims to learn the mapping from eye movements to the ideal weigths. The mapping is learned, within each search topic, from the eye measurements on a word to its associated ideal weight value. The learned mapping can then be used to infer a new weight vector for new documents and new interests; the weight vector is in turn used in a classifier on new unseen documents. It is the inferred implicit "query."

In addition to a linear regressor, we use a non-linear kernel Partial Least Squares (KPLS) (Rosipal & Trejo, 2001). In this method the features are projected non-linearly to a new subspace where a standard least squares regression is performed. A set of alternative criteria for selecting the projection directions have been introduced(Dhanjal, Gunn, & Shawe-Taylor, 2006); in this task we choose a sparse combination of features having maximal covariance. We will make experiments on several projection subspace dimensionalities.

We use a Gaussian kernel with the sparse-KPLS where the width parameter σ for the kernels is optimized, per search topic, using 10-fold cross validation on the training data.

5 EXPERIMENTAL SET-UP

We apply the Term Frequency Inverse Document Frequency (TFIDF) (Salton & McGill, 1983) on the documents in order to increase the weighting of terms that occur frequently within a document but infrequently across the corpus. If our corpus is given by D and dictionary by T, then the TFIDF for term $t \in T$ in document $d \in D$ is given by

$$\mathrm{TFIDF}(d,t) = n_{dt} \log \frac{|D|}{|\{\delta \in D \mid n_{\delta t} > 0\}|},$$

where n_{dt} denotes the number of occurrences of term t in document d. The TFIDF representation is used throughout the experiments.

5.1 EYE MOVEMENT AND TEXT FEATURE MODELS

Our main goal is to study whether it is possible to discriminate between the categories using a weight vector inferred only from the corresponding eye movement and text features. The set-up for this task is as follows; for each search topic we do the following:

- We compute the "ideal weights" by optimizing a SVM to classify the documents belonging to this search topic vs. the others.
- Using the eye and text features described in Section 3, we train a regressor from the feature vectors to the corresponding ideal term weights (to clarify: a single regressor having a scalar output is applied to all terms). To make sure that this regressor is universal and will not utilize any category-specific information, the training data comes from all the other categories instead of the current left-out category. The regressor is either a standard linear least squares regressor or the non-linear sparse-KPLS described in Section 4.2. We normalise the regressor weights in the 2-norm.
- Using the learnt regressor we compute a weight vector for the left out search topic. Here the inputs are eye movement features of documents in the left out category; feature vectors for reoccurring words are averaged over users. Zero is assigned to terms that had no fixations. We refer to the inferred weight vector classifier as W_i when the regressor is linear, and as $W_i(x)$ with the nonlinear regression, as described in subsection 4.2. Here x is the number of feature directions used in the KPLS and the subindex i refers to the index of the topic emphasizing the fact that there is a separate classifier for each topic.

Figure 2 shows a graphical representation of the predicted weights for one document and one search topic.

Here, as well as in the Section 5.2 when combining the eye movements with explicit feedback, we ignore the eye movements on terms that do not belong to the dictionary.

5.2 COMBINING IMPLICIT AND EXPLICIT FEEDBACK

To construct a relevance predictor for the explicit feedback scenario we train an SVM classifier. The inputs for the SVM are the TFIDF vectors of the documents in one search topic and the outputs are the explicit relevance labels for the same documents given by the users. We refer to this classifier as SVM_i.

The SVM-$2\mathrm{K}_i$ classifier that uses both explicit relevance inputs and implicit eye movements feedback is trained using the same documents. We assume that the eye movements provide a second representation of the documents and we concatenate the eye movement and text features for words occurring in that document. We use the TFIDF vectors and the concatenated eye movement features as the two inputs for an SVM-2K classifier.

Note that for this case the eye movement features are not associated with individual words unlike in the earlier experiment. Furthermore, in the testing phase there are no eye movements features available so only the text is taken into account (they are an average for word reoccurrence).

6 RESULTS

Evaluation criteria. The SVM predictor is used to rank the new documents in the order of expected relevance. Specifically, we compute the measure of performance by first ranking the 244 unseen documents in the test set according to their predicted relevance to the user. The document predicted the most relevant has the rank one, and the document predicted the least relevant has a ranking of 244. The *average precision* is given by

$$PREC = \frac{1}{R} \sum_{i=1}^{R} \frac{i}{r_i}, \qquad (1)$$

where R is the number of positive examples in the test set of 244 documents, and r_i are rankings of the positive examples, ordered such that $r_i < r_{i+1}$. If the classifier is working optimally, that is, we predict the positive examples to have rankings $1, \ldots, R$, i.e., $r_i = i$, we obtain a precision of 100%. Notice that for conciseness and in a slight abuse of terminology, in the

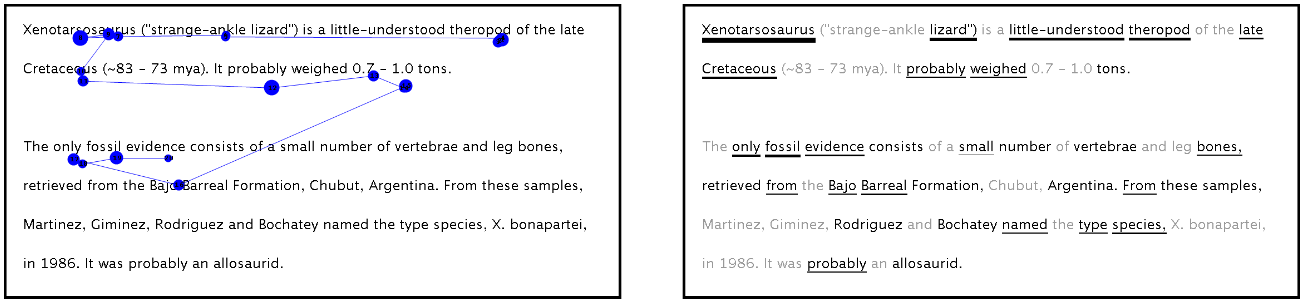

Figure 2: Sample plot of saccades (lines) and fixations (dots) on a document (on the left) and term weights inferred from eye movements on all documents in the *Dinosaurs* category (on the right). The magnitude of the inferred weight is shown by the thickness of the underlining. The words which do not appear in the dictionary are shown in light grey.

following we will use *precision* to refer to the average precision defined above.

Baseline models. To show that including the eye movements in the model really is beneficial in the document relevancy prediction task we compare the result of the model using all 26 eye movement and text features to a model that uses only the 4 text features. Both models are trained as was described in Section 5.1. The only difference is in the feature sets. In Table 2 the text features only model is denoted by $W_{\text{text}}(4)$. The average precision for this model is 29.63%.

The expected precision of a uniformly random ranking is 3.69% for 4 positive examples out of 244 documents (the *Natural disasters* category) and 6.10% for 10 positive examples (all other categories). These results are significantly worse than our other results because there is an imbalance in the proportion of positive examples in the training (about half are positive) and test sets (only 10 out of 244 are positive) which the random model does not take into account. We have controlled this bias by comparing two models $W_{\text{text}}(4)$ and $W_i(26)$ (see below) which differ only in that in latter also the eye movements have been taken into account

The upper expected limit of performance is given by the *ideal weights*, denoted by SVM in the results in Table 2.

Models with combined features. The non-linear regression model $W_i(26)$ uses both the eye movement and the term features and the number of projection directions in the KPLS regression equals the number of features. It has average precision of 39.82%. The result is significantly better than that of the text features only model ($P < 0.01$, Fisher Sign Test). This is quite a strong result considering the complexity of the task.

We tested also two other non-linear models, labelled $W_i(39)$ and $W_i(52)$, with the number of projection directions exceeding the number of features. They have similar overall performance. The linear regression model W_i has a bit lower precision on average. The topicwise results are shown in Table 2.

It is interesting to observe that the some search topics achieve a higher precision with the linear regression model than with non-linear one. Despite these results it is apparent that the non-linear approach outperforms the linear one across all selections of the number of feature directions.

It is striking that the eye movement models perform worse than the text features only baseline model in some categories. One possible reason for that is that some users have read through most of some documents instead of just finding enough evidence to judge the relevance, perhaps because they were interested in the topic. This kind of reading behaviour would not emphasise the interesting words and would make it impossible to learn the regressor.

Eye movements combined with explicit relevance feedback and text content. Our initial assumption was that combining eye movements with the explicit relevancy feedback improves overall performance. Comparing SVM$_i$ and SVM-2K$_i$ results in the Table 2 shows that this is not true for all search topics. Nevertheless, the overall precision is improved by combining the two sources of information.

7 DISCUSSION

We addressed the extremely hard task of constructing a query in an information retrieval task, given neither an explicit query nor explicit relevance feedback. Only eye movement measurements for a small set of viewed snippets, and the text content of the snippets

Table 2: The precision for various predictors and search topics, in percent. Larger precision is better. The baseline models are a text-feature-only model and an SVM constructed directly of the documents to classify them into 25 topic classes. The baseline models provide expected lower and upper limits for the supposed performance of the predictors. The largest precision for each search topic and class of predictors is shown in boldface. The $W_i(26)$ model outperforms the $W_{\text{text}}(4)$ baseline model ($P < 0.01$, Fisher Sign Test).

	Baseline		Eye movements				Expl. feedb.	Impl.&Expl. feedback
	$W_{\text{text}}(4)$	SVM	W_i	$W_i(26)$	$W_i(39)$	$W_i(52)$	SVM$_i$	SVM-2K$_i$
Astronomy	19.33	63.66	**24.18**	17.75	16.95	18.35	39.57	**40.11**
Ball games	64.50	100.00	**85.91**	66.57	59.88	57.26	75.33	**86.01**
Cities	15.38	96.22	**25.14**	20.91	19.35	21.31	69.83	**80.53**
Court systems	47.70	85.67	47.38	**53.01**	49.05	46.21	**62.72**	59.83
Dinosaurs	38.42	100.00	53.42	68.49	**71.16**	63.30	**95.73**	94.30
Education	26.53	96.69	30.60	44.98	**54.32**	38.98	50.33	**56.25**
Elections	42.26	75.67	42.09	43.30	44.70	**44.93**	**72.87**	67.52
Family	23.03	83.54	22.22	**25.83**	25.75	21.19	70.22	**71.81**
Film	10.51	81.02	23.28	52.20	**53.72**	45.28	**54.51**	54.08
Government	28.25	68.80	32.88	**35.67**	35.35	35.24	**32.75**	26.92
Internet	5.24	67.10	8.84	7.37	7.91	**9.36**	35.58	**39.35**
Languages	49.24	96.52	55.27	81.84	74.89	**81.98**	89.74	**93.51**
Literature	12.61	56.80	14.32	16.59	**22.69**	15.70	18.24	**26.84**
Music	8.92	82.67	**16.32**	11.89	12.26	11.19	60.21	**74.03**
Natural disasters	73.33	100.00	83.04	85.42	**88.75**	**88.75**	**100.00**	**100.00**
Olympics	27.67	98.09	39.80	39.13	**42.60**	37.27	92.63	**97.69**
Optical devices	16.83	81.69	18.21	**18.82**	15.90	18.13	56.08	**64.44**
Postal system	25.56	99.09	20.44	**29.39**	20.72	20.66	76.30	**81.66**
Printing	48.89	100.00	55.12	60.44	**63.68**	62.39	63.24	**68.01**
Sculpture	7.03	86.35	19.52	**25.01**	20.70	24.94	60.17	**62.44**
Space exploration	45.74	94.07	72.16	75.03	72.46	**75.60**	65.08	**67.41**
Speeches	27.20	84.80	37.30	**42.26**	37.40	37.09	**75.31**	70.29
Television	39.74	88.79	45.20	40.29	**49.83**	42.68	**36.68**	34.61
Transportation	16.86	70.22	13.07	**13.12**	11.57	11.68	**44.02**	41.59
Writing systems	19.91	95.56	28.63	20.26	**31.12**	25.31	46.76	**50.28**
Average	29.63	86.12	36.57	39.82	40.11	38.19	61.76	64.38

was available. This is a prototype of a task where the intent or interests of the user are inferred from implicit feedback signals, and used to anticipate the users actions.

We were able to learn a "universal predictor of relevance predictors" from a collected database of queries, their relevant and irrelevant documents, and the corresponding eye movements. The predictions performed better than a simple model which utilized only the textual content of the documents on new queries. There is ample room for improvement in the prediction percentages; our best eye movements model gives mean precision of 40.11% as opposed to 29.63% of the text only model. Nonetheless, the feasibility study was successful; the conclusion is that the eye movement help in predicting relevancy of a document.

We further experimented with a model where the textual content of the documents and explicit relevance feedback given by the user (whether or not the user thinks the document is relevant to the search topic) were taken into account. As expected, the pure explicit feedback improved the precision significantly to 61.76%. Our results show that also in this scenario, taking the eye movements into account we can further improve the precision by a couple of percentage points.

We conclude that in constructing a query, eye movements provide a useful implicit feedback channel. As expected, the feedback obtained from the eye movements is less informative than relevance feedback typed in by the user, but nonetheless this implicit feedback can be exploited. In practical applications all available feedback channels, in addition to the eye movements, should of course be utilized; the practical implication of this study is that if eye movement data is cheaply available it might be a good idea to include it as well.

In this work, we used relatively standard state-of-the-art machine learning methods. An obvious direction for future research would be to develop a tailored method to obtain a query from eye movements, optimised as one single model. Another obvious extension is to investigate how well the relevance predictors learned in relatively controlled experimental setups generalize to more practical situations. The alternative is to learn the predictors on-line in a practical application.

Acknowledgements

AA and SK belong to Adaptive Informatics Research Centre, a centre of excellence of the Academy of Finland. The authors thank Wray Buntine from the Complex Systems Computation Group at University of Helsinki for providing the Wikipedia data and Craig Saunders and Steve Gunn from the ISIS Research Group at University of Southampton for many fruitful discussions. This work was supported in part by the IST Programme of the European Community, under the PASCAL Network of Excellence, IST-2002-506778. This publication only reflects the authors views.

All rights are reserved because of other commitments.

References

Dhanjal, C., Gunn, S. R., & Shawe-Taylor, J. (2006). Sparse feature extraction using generalised partial least squares. In *Proceedings of the IEEE international workshop on machine learning for signal processing* (pp. 27–32).

Farquhar, J. D. R., Hardoon, D. R., Meng, H., Shawe-Taylor, J., & Szedmak, S. (2006). Two view learning: Svm-2k, theory and practice. In Y. Weiss, B. Schölkopf, & J. Platt (Eds.), *Advances in neural information processing systems 18* (pp. 355–362). Cambridge, MA: MIT Press.

Hardoon, D. R., Szedmak, S. R., & Shawe-taylor, J. R. (2004). Canonical correlation analysis: An overview with application to learning methods. *Neural Computation*, *16*(12), 2639–2664.

Joachims, T., Granka, L., Pan, B., Hembrooke, H., & Gay, G. (2005). Accurately interpreting clickthrough data as implicit feedback. In *Proceedings of the 28th annual international ACM SIGIR conference on research and development in information retrieval* (pp. 154–161).

Kelly, D., & Teevan, J. (2003). Implicit feedback for inferring user preference: a bibliography. In *ACM SIGIR forum* (pp. 18–28).

Maglio, P. P., Barrett, R., Campbell, C. S., & Selker, T. (2000). SUITOR: An attentive information system. In *IUI-2000: International conference on intelligent user interfaces* (pp. 169–176).

Maglio, P. P., & Campbell, C. S. (2003). Attentive agents. *Communications of the ACM*, *46*(3), 47–51.

Puolamäki, K., & Kaski, S. (Eds.). (2006, May). *Proceedings of the NIPS 2005 workshop on machine learning for implicit feedback and user modeling.* Otaniemi, Finland. (http://www.cis.hut.fi/inips2005/)

Puolamäki, K., Salojärvi, J., Savia, E., Simola, J., & Kaski, S. (2005). Combining eye movements and collaborative filtering for proactive information retrieval. In G. Marchionini, A. Moffat, J. Tait, R. Baeza-Yates, & N. Ziviani (Eds.), *Proceedings of SIGIR 2005, twenty-eighth annual international ACM SIGIR conference on research and development in information retrieval* (pp. 146–153). New York, NY: ACM.

Rosipal, R., & Trejo, L. J. (2001). Kernel partial least squares regression in reproducing kernel hilbert space. *Journal of Machine Learning Research*, *2*, 97–123.

Salojärvi, J., Kojo, I., Simola, J., & Kaski, S. (2003, September). Can relevance be inferred from eye movements in information retrieval? In *Proceedings of WSOM'03, workshop on self-organizing maps* (pp. 261–266). Hibikino, Kitakyushu, Japan: Kyushu Institute of Technology.

Salojärvi, J., Puolamäki, K., & Kaski, S. (2005). Implicit relevance feedback from eye movements. In W. Duch, J. Kacprzyk, E. Oja, & S. Zadrozny (Eds.), *Artificial neural networks: Biological inspirations – ICANN 2005* (pp. 513–518). Berlin, Germany: Springer-Verlag.

Salojärvi, J., Puolamäki, K., Simola, J., Kovanen, L., Kojo, I., & Kaski, S. (2005, March). *Inferring relevance from eye movements: Feature extraction* (Tech. Rep. No. A82). Helsinki University of Technology, Publications in Computer and Information Science. (http://www.cis.hut.fi/eyechallenge2005/)

Salton, G., & McGill, M. J. (1983). *Introduction to modern information retrieval.* New York, NY, USA: McGraw-Hill, Inc.

Turpin, A., & Scholer, F. (2006). User performance versus precision measures for simple search tasks. In *Proceedings of the 29th annual international ACM SIGIR conference on research and development in information retrieval* (pp. 11–18).