Interactive Intent Modeling from Multiple Feedback Domains

Pedram Daee[1], Joel Pyykkö[2], Dorota Głowacka[2], and Samuel Kaski[1]
Helsinki Institute for Information Technology HIIT
[1]Aalto University, Department of Computer Science
[2]University of Helsinki, Department of Computer Science
firstname.lastname@hiit.fi

ABSTRACT
In exploratory search, the user starts with an uncertain information need and provides relevance feedback to the system's suggestions to direct the search. The search system learns the user intent based on this feedback and employs it to recommend novel results. However, the amount of user feedback is very limited compared to the size of the information space to be explored. To tackle this problem, we take into account user feedback on both the retrieved items (documents) and their features (keywords). In order to combine feedback from multiple domains, we introduce a *coupled multi-armed bandits* algorithm, which employs a probabilistic model of the relationship between the domains. Simulation results show that with multi-domain feedback, the search system can find the relevant items in fewer iterations than with only one domain. A preliminary user study indicates improvement in user satisfaction and quality of retrieved information.

Author Keywords
Exploratory search; intent modeling; multi-armed bandits; relevance feedback; probabilistic user models.

ACM Classification Keywords
H.3.3 Information Search and Retrieval: Relevance feedback; I.2.6 Learning.

INTRODUCTION
The dominant information retrieval paradigm relies on the user's ability to form a precise query, which is difficult at least in the about 50% of search sessions where the user is uncertain about her information need [17]. Furthermore, the information need can shape throughout the search session. For instance, when the user prepares for writing a summary about a particular topic, the search process typically takes several iterations in which the user directs the search by tuning the initial query and the initial search intent, after observing the results. This important type of a search scenario is called *exploratory search* [10, 12, 18].

IUI'16, March 07–10, 2016, Sonoma, CA, USA.

DOI: http://dx.doi.org/10.1145/2856767.2856803

There is a wide variety of qualitative definitions of exploratory search [19]. Marchionini [12] illustrated exploratory and lookup tasks as an overlapping cloud and suggested that lookup tasks are embedded in exploratory tasks and vice versa. The problem context that motivates the search process is typically characterized in definitions of exploratory search [18]. Imprecise task requirements or open-ended search goals are the two primary attributes often used to define exploratory search with respect to the problem context [11]. The exploratory search process is considered to be cognitively complex with the information seeker being uncertain about the search process [18].

To help the user in exploratory search, her interactions with the system can be employed to infer her search intent. This is challenging for two reasons. First, active interaction is required from the user; however, users are often not willing to invest in actively giving feedback to search systems. Second, even if the interface was appealing enough, the user can only provide a limited amount of feedback, which makes user intent modelling challenging. In this paper we make it easier for the user to give feedback in exploratory search, by allowing feedback on multiple domains, in this case keywords and documents. For this we formulate the user intent modelling task as a new coupled multi-armed bandit problem, instead of using only one multi-armed bandit for one modality as in earlier approaches [14].

The rest of the paper is organized as follows: First we model exploratory search as a learning problem with limited feedback. Next, we propose the *coupled multi-armed bandits* algorithm that employs Thompson sampling with a novel probabilistic user model. We conclude the paper by evaluating the proposed method in a simulation scenario and a user study.

PROPOSED APPROACH

Problem Setting
Let D be a set of documents in a corpus and K be a set of keywords extracted from these documents. For each user, it is assumed that the relevance of each document $d \in D$ is an unknown distribution over $[0,1]$. This distribution encodes the uncertainty of the user about the relevance of each item, which is a key element of exploratory search. The expected

relevance of d for the user is denoted by $E_D[d]$. The document $d \in D$ is more relevant than $d' \in D$, if $E_D[d'] < E_D[d]$. Similarly, it is assumed that the relevance of each keyword $k \in K$ is an unknown distribution over [0,1] with its expected relevance denoted by $E_K[k]$. We call this set of distributions the user intent model. The expected relevances of keywords and documents are connected through a model of the data that defines how keywords belong to documents.

In each session, a user with a fixed but unknown intent model arrives. Each session consists of N iterations. In each iteration, the user provides input based on her intent model as relevance values to keywords and documents, and the algorithm provides a set of new documents and keywords. It is assumed that user feedback consists of samples from the relevance distributions.

The retrieval system should look for the most relevant document $d^* = \arg\max_{d \in D} E_D[d]$ and present it to the user. To solve this maximization the system needs to explore the document space to estimate the expected relevance of documents based on user feedback. At the same time it should exploit the estimates to show relevant documents as early as possible. This kind of a black box optimization problem, where the objective function is unknown and expensive to sample, has been studied in multi-armed bandit [6] and Bayesian optimization [4, 16] literature. A natural performance criterion for these problems is regret, which is the loss due to not presenting the most relevant documents to the user. The cumulative regret after receiving feedback for a set of documents n_D is $cum\_regret = |n_D| E_D[d^*] - \sum_{d \in n_D} E_D[d]$. The goal is to minimize the cumulative regret, which is equivalent to maximizing the sum of expected relevance feedback on documents in n_D. However, since the expected relevance of documents is hidden to the system, this measure cannot be calculated in practice.

In practice, an exploratory search system is successful if it presents to the user items, e.g. documents, that the user finds interesting. How we measure this "interest" is through maximizing the average number of clicks (or other types of positive relevance feedback) on documents in N iterations [15]. Since it is reasonable to assume that the user provides positive feedback mostly on relevant documents, this user behavior model also minimizes the cumulative regret defined from the theory point of view.

By modeling the problem as regret minimization, there is little hope to achieve a reasonable result by only considering the limited feedback on documents. In exploratory search, it is more convenient for the user to express the abstract understanding of her needs in terms of higher-level information, such as a set of keywords. In this paper we take advantage of both feedback on documents and on keywords in order to improve the regret of an exploratory search system. We tackle the problem by modeling it as coupled multi-armed bandits where it is possible to provide feedback both to the arms and the features defining the arms. In our exploratory search problem, the arms are documents and the features are keywords defining the documents.

Connecting Documents and Keywords

We assume there exists a document-keyword matrix M defined as

$$M = \begin{bmatrix} P(k_1|d_1) & \cdots & P(k_{|K|}|d_1) \\ \vdots & \ddots & \vdots \\ P(k_1|d_{|D|}) & \cdots & P(k_{|K|}|d_{|D|}) \end{bmatrix}_{|D| \times |K|}$$

where $P(k_i|d_j)$ specifies the likelihood of document d_j generating keyword k_i. This matrix is generated from the data model that expresses how keywords and the documents are related. An example, which we use, is to consider documents as bags of words and to use normalized *tf-idf* representations of documents. We make the simplifying assumption that the connection between expected relevance of a document and keywords is as follows:

$$E_D[d_j] = \sum_{i=1}^{|K|} E_K[k_i] P(k_i|d_j)$$

With the compact notation $E_D[D] = \left[E_D[d_1], \ldots, E_D[d_{|D|}]\right]^T$, and analogously for keywords, this becomes

$$E_D[D] = M E_K[K] \qquad (1)$$

Note that we made the simplifying assumptions only for the expected relevances; the shapes of the relevance distributions can be different.

Coupled Bayesian Bandits

The user provides relevance feedback to both documents and keywords. The relevance (reward) distributions for document d and keyword k are denoted by $r_d \sim f^D(.|x_d, \theta)$ and $r_k \sim f^K(.|x_k, \theta)$, respectively. The x_d and x_k are the feature (context) vectors associated with d and k. The fixed but unknown parameter θ defines the shared link between these two relevance distributions. After the user has interacted with a set of documents n_D and a set of keywords n_K, we can write the posterior at time $t = |n_D| + |n_K|$ as

$$\pi_t(\theta) \propto \pi_0(\theta) \prod_{d \in n_D} f^D(r_d|x_d, \theta) \prod_{k \in n_K} f^K(r_k|x_k, \theta), \qquad (2)$$

where $\pi_0(\theta)$ is the prior distribution for θ. In order to apply Bayesian bandit methods [1, 9] it is only necessary to be able to perform the following two steps: draw a sample from the posterior at time t, and after that score all the documents and keywords by $E_D[r_d|x_d, \theta^p]$ and $E_K[r_k|x_k, \theta^p]$.

These two steps are the minimum requirements for employing the Thompson sampling algorithm for bandits [1, 7, 9]. In Thompson sampling, the exploration and exploitation are controlled indirectly by the uncertainty in the posterior. We can easily draw a sample from the posterior by using any sampling method, after specifying the shared parameter θ that connects the relevance distributions, and the feature vectors x_d and x_k. These parameters are defined by the user model.

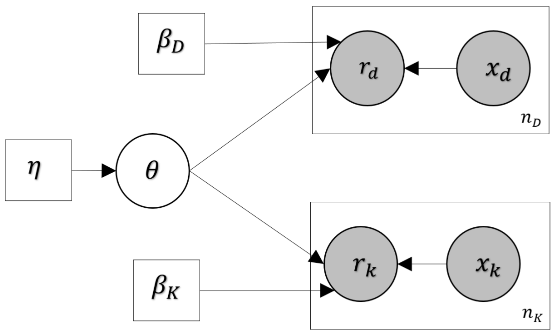

Figure 1. Probabilistic model for user feedback on documents and keywords.

Probabilistic User Model

We propose a simple model for received relevance feedback on documents and keywords. Since the amount of feedback from the user is limited, we need to impose a structure on the expected relevance of items to be able to generalize well. We assume that the expected relevance of keywords is linearly related to their feature vectors by the unknown weight vector θ, i.e. $E_K[K] = M^T\theta$. Based on this linearity assumption and our previous assumption in equation (1), we have $E_D[D] = ME_K[K] = MM^T\theta$. Using this feature transformation, we only need to estimate one set of unknown weights to specify expected relevance of both documents and keywords. Considering Gaussian distributions for relevance, we propose the following model (plate diagram in Figure 1):

$$f^K(r_k|x_k,\theta,\beta_K) = N(r_k; x_k^T\theta, \beta_K^2)$$

$$f^D(r_d|x_d,\theta,\beta_D) = N(r_d; x_d^T\theta, \beta_D^2)$$

$$\pi_0(\theta) = N(\theta; 0, \eta^2\mathrm{I})$$

Here, x_k is the k^{th} column of M and x_d is the d^{th} column of MM^T; they define feature vectors for keyword k and document d, respectively. Since all the distributions are Gaussian, the posterior is also a Gaussian distribution, with:

$$\pi_t(\theta) = N(\theta; \mu_t, \Sigma_t)$$

$$\Sigma_t^{-1} = \beta_D^{-2}X_{n_D}^T X_{n_D} + \beta_K^{-2}X_{n_K}^T X_{n_K} + \eta^{-2}\mathrm{I} \quad (3)$$

$$\mu_t = \Sigma_t\left(\beta_D^{-2}X_{n_D}^T R_{n_D} + \beta_K^{-2}X_{n_K}^T R_{n_K}\right)$$

where X_{n_D} is a $|n_D| \times D$ design matrix containing feature vectors for the observed documents in the set n_D, and R_{n_D} is a $|n_D| \times 1$ matrix of the corresponding observed relevance values. With the same logic, X_{n_K} is a $|n_K| \times D$ design matrix containing feature vectors for the observed keywords in the set n_K, and R_{n_K} is a $|n_K| \times 1$ matrix of the corresponding observed relevance values. The computational complexity of (3) comes from the covariance matrix inversion.

The coupled multi-armed bandits algorithm employing Thompson sampling for controlling exploration-exploitation tradeoff is as follows:

ALGORITHM 1: Coupled multi-armed bandits

At time step t
1. draw $\theta^p \sim \pi_t(\theta)$ from (3)
2. for document bandit: select $d^+ = \arg\max_{d\in D} x_d^T\theta^p$
3. for keyword bandit: select $k^+ = \arg\max_{k\in K} x_k^T\theta^p$
4. update the posterior (3) based on user feedback and observed feature vectors

SIMULATION

In each exploratory search session, the simulator considers a small set of documents and keywords as targets, and assumes their expected relevance to be 1. The simulated user employs Latent Semantic Indexing (LSI) and cosine similarity measure to calculate similarities to the targets for all documents and keywords. According to these similarities, an expected relevance value in [0,1] is assigned to each document and keyword, defining $E_D[D]$ and $E_K[K]$.

In each iteration, the algorithm presents five documents and five keywords to the user (by repeating steps 2 and 3 of Algorithm 1). We assume that the simulated user selects document d as relevant with probability equal to relevance value $r_d \sim N(E_D[d], \beta_D^2)$, and provides relevance feedback $r_k \sim N(E_K[k], \beta_K^2)$ for keyword k with probability proportional to $|r_k - 0.5|$. The motivation is that users usually give feedback to keywords that are highly relevant or irrelevant. We used a fixed pool of 750 computer science arXiv articles, with only 50 of them having high expected relevance in the user intent model. The model parameters were set to $\beta_D = \beta_k = 0.3$ and $\eta = 0.5$. We compared our method with variants that only consider feedback on keywords or documents to update the posterior (3).

Based on the simulation result in Figure 2, we can conclude that considering feedback on both documents and keywords will significantly increase the number of relevant items that the user can find with the same number of iterations.

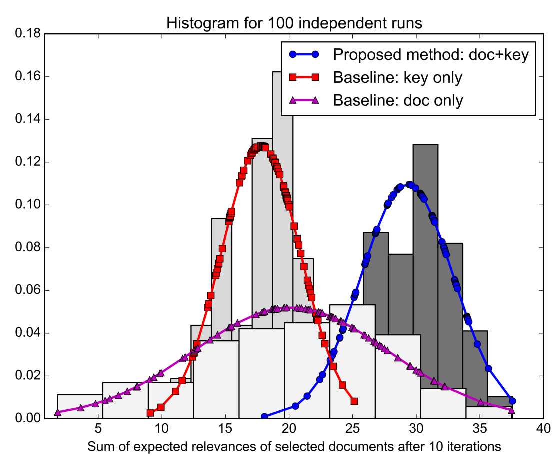

Figure 2. Histogram for 100 independent runs of the methods. The sum of expected relevances of selected documents after 10 iterations is significantly higher when both types of feedback (doc+key) are considered.

USER EXPERIMENTS AND RESULTS

We implemented the proposed method in an existing exploratory search system SciNet [8]. SciNet uses the LinRel algorithm [2] to learn the user intent model based on feedback on keywords visualized on a radar-like interface. In our implementation, we replaced LinRel with Algorithm 1. Furthermore, user interactions with documents, i.e. clicks or bookmarks, are considered as relevance feedback on documents, in addition to feedback on keywords. The computational complexity of both algorithms is the same.

We conducted a preliminary user study on 10 university students and researchers. All the participants were fluent in English and had some background in computer science or a related field. Each participant performed two exploratory tasks in which they had to do a literature survey on a topic and answer three questions in a fixed amount of time. The topics were reinforcement learning and neural networks. The participants reported their knowledge of the topics on 1-5 Likert scale (with 2.2 on average). The user interface and real time performance of the systems were identical, and the participants were naïve about which exploratory search engine was used for each task. In each iteration, five documents were shown to the user and the user could bookmark the relevant ones. All user interactions were logged by the system including the typed query, documents and keywords presented to the user, and documents and keywords that the user interacted with. After each task the participants answered a questionnaire containing SUS [5] and a short version of ResQue [13] using 1-5 Likert scale, where 1 indicates “strongly disagree”. A short interview with each participant was conducted after the tasks.

Answers to all the tasks and the bookmarked documents were rated by experts in a double-blind assessment using a scale from 0 (no answer) to 5 (perfect answer). All the shown documents were assessed on a binary scale based on their relevance to the search topic. The inter-rater agreement between the experts showed that the rankings overlapped more than 80%. One of the users was excluded from the analysis because he did not bookmark any article.

Table 1 summarises different performance measures for the proposed algorithm and the baseline system. The percentage of all the shown articles that were labelled as relevant by the experts was calculated as a measure of the quality of the shown information. For the proposed system this value was 5 percent better compared to the baseline. However, the difference was not statistically significant.

Performance measure	Proposed	Baseline
Average SUS score	**75**	67.2
Average ResQue score	**54.7**	52.3
% of relevant documents	**84.6**	79.1
User task performance	3.45	3.45

Table 1. Performance measures for the proposed and baseline systems. The better value on each row is shown in bold.

The proposed algorithm had better SUS and ResQue scores compared to the baseline. However, due to the small sample size these differences were not statistically significant. It should be noted that most of the questions in SUS and ResQue target areas such as user interface, which was the same in both systems. There are two questions in ResQue that measure the *novelty* and *diversity* of search results [13]: *The recommender system helped me discover new items* and *The items recommended to me are diverse*. In both questions the proposed method scored higher (4.1 and 3.7) against the baseline system (3.5 and 3.3).

User task performance was measured by averaging the expert assessments of the answers for the three questions in each task. The users were hard pressed to gather a report in time, which was also evident in their answers. In these reports both systems achieved the same average performance.

In the interviews, 6 out of 9 users reported higher satisfaction with the proposed system (more diversity and better results), one reported that he could not tell the difference, and two said that the baseline was better.

Overall, the preliminary results indicate that the proposed system gave the users a more satisfying image of the topic they were exploring. Furthermore, the usability of the total system has also improved.

DISCUSSION AND CONCLUSION

In this paper we introduced the *coupled multi-armed bandits* algorithm as the exploratory search method that employs the user feedback on both the retrieved items and their features. Our approach is based on two main ideas. First, we model user behavior as a generative probabilistic model. Second, we couple different sources of feedback in a unified model. Our simulation results and preliminary user study indicate that considering these two sources of feedback can improve the performance and quality of the exploratory search.

From the practical point of view, our algorithm provides the opportunity to exploit several types of relevance feedback that are available for documents, in addition to relevance feedback available for individual keywords. For example in [3] it was shown that it is possible to detect implicit relevance feedback from physiological signals such as electrodermal activity and facial electromyography on documents. We believe that this is an important step for the future of exploratory search applications since an increasing amount of research studies the feasibility of performing information retrieval based on novel types of relevance signals both on keywords and on documents.

ACKNOWLEDGMENTS

This work has been partly supported by the Academy of Finland (Finnish Centre of Excellence in Computational Inference Research COIN), Re:Know funded by TEKES, and MindSee (FP7–ICT; Grant Agreement #611570).

REFERENCES

1. Agrawal, S., and Goyal, N. Thompson Sampling for Contextual Bandits with Linear Payoffs. In *Proc. of ICML,* (2013), 127-135.
2. Auer, P. Using confidence bounds for exploitation-exploration trade-offs. *The Journal of Machine Learning Research*, (2003), 397-422.
3. Barral, O., Eugster, M. J., Ruotsalo, T., Spapé, M. M., Kosunen, I., Ravaja, N., ... and Jacucci, G. Exploring Peripheral Physiology as a Predictor of Perceived Relevance in Information Retrieval. In *Proc. of IUI,* ACM (2015), 389-399.
4. Brochu, E., Cora, V. M., and De Freitas, N. A tutorial on Bayesian optimization of expensive cost functions, with application to active user modeling and hierarchical reinforcement learning. *arXiv preprint arXiv:1012.2599*, 2010.
5. Brooke, J. SUS-A quick and dirty usability scale. *Usability evaluation in industry 189*, (1996), 4-7.
6. Bubeck, S., and Cesa-Bianchi, N. Regret analysis of stochastic and nonstochastic multi-armed bandit problems. *Foundations and Trends in Machine Learning*, (2012), 5(1), 1-122.
7. Chapelle, O., and Li, L. An empirical evaluation of thompson sampling. In *Advances in Neural Information Processing Systems,* (2011), 2249-2257.
8. Glowacka, D., Ruotsalo, T., Konuyshkova, K., Kaski, S., and Jacucci, G. Directing exploratory search: Reinforcement learning from user interactions with keywords. In *Proc. of IUI,* ACM (2013), 117-128.
9. Hoffman, M. D., Shahriari, B., and de Freitas, N. On correlation and budget constraints in model-based bandit optimization with application to automatic machine learning. In *Proc. of AISTATS,* (2014), 365-374.
10. Kammerer, Y., Nairn, R., Pirolli, P. and Chi, E.H. Signpost from the masses: learning effects in an exploratory social tag search browser. In *Proc. of the SIGCHI Conference on Human Factors in Computing Systems*, ACM (2009), 625-634.
11. Kim, J. Describing and predicting information-seeking behavior on the web. *Journal of the American Society for Information Science and Technology*, (2009), 679–693.
12. Marchionini, G. Exploratory search: from finding to understanding. *Communications of the ACM* (2006), 41–46.
13. Pu, P., Chen, L., and Hu, R. A user-centric evaluation framework for recommender systems. In *Proc. of RecSys,* ACM (2011), 157-164.
14. Ruotsalo, T., Jacucci, G., Myllymäki, P. and Kaski, S. Interactive intent modeling: Information discovery beyond search. *Communications of the ACM*, (2014), 58(1), 86-92.
15. Slivkins, A., Radlinski, F., and Gollapudi, S. Ranked bandits in metric spaces: learning diverse rankings over large document collections. *The Journal of Machine Learning Research*, (2013), 14(1), 399-436.
16. Srinivas, N., Krause, A., Kakade, S. M., and Seeger, M. Gaussian process optimization in the bandit setting: No regret and experimental design. In *Proc. of ICML,* (2010).
17. Teevan, J., Alvarado, C., Ackerman, M. S., and Karger, D. R. The perfect search engine is not enough: a study of orienteering behavior in directed search. In *Proc. of CHI,* ACM (2004), 415-422.
18. White, R. W., and Roth, R. A. Exploratory search: Beyond the query-response paradigm. *Synthesis Lectures on Information Concepts, Retrieval, and Services*, (2009), 1-98.
19. Wildemuth, B. M., and Freund, L. Assigning search tasks designed to elicit exploratory search behaviors. In *Proc. of HCIR*. ACM (2012), 1–10