

Effects of Foraging in Personalized Content-based Image Recommendation

Amit Kumar Jaiswal
University of Bedfordshire
Luton, United Kingdom
amitkumar.jaiswal@beds.ac.uk

Haiming Liu
University of Bedfordshire
Luton, United Kingdom
haiming.liu@beds.ac.uk

Ingo Frommholz
University of Bedfordshire
Luton, United Kingdom
ingo.frommholz@beds.ac.uk

ABSTRACT

A major challenge of recommender systems is to help users locating interesting items. Personalized recommender systems have become very popular as they attempt to predetermine the needs of users and provide them with recommendations to personalize their navigation. However, few studies have addressed the question of what drives the users' attention to specific content within the collection and what influences the selection of interesting items. To this end, we employ the lens of Information Foraging Theory (IFT) to image recommendation to demonstrate how the user could utilize visual bookmarks to locate interesting images. We also found that the recommended image collection with visual bookmarks (cues) leads to a stronger scent. In this paper, we investigate a personalized content-based image recommendation system to understand what expropriates user attention by reinforcing visual attention based on IFT. Our evaluation is based on the Pinterest image collection.

CCS CONCEPTS

• **Information systems** → **Personalization; Recommender systems**; • **Human-centered computing** → *Web-based interaction*;

KEYWORDS

Information Foraging Theory, Information Retrieval, Recommendation System

ACM Reference Format:

Amit Kumar Jaiswal, Haiming Liu, and Ingo Frommholz. 2019. Effects of Foraging in Personalized Content-based Image Recommendation. In *Proceedings of EARS '19: The 2nd International Workshop on Explainable Recommendation and Search (EARS 2019) (EARS '19)*. ACM, Paris, France, 5 pages. <https://doi.org/10.1145/xxxxxx.xxxxxxx>

1 INTRODUCTION

Searching the Web is an important part of many people's everyday life. Retrieved online content is usually the outcome of generic user searches include textual documents, images, etc. As of now, the general searching method users employ is based on keywords, which is supported by almost every commercial search engine. Often users are also pointed to information by means of recommendation, which can for instance be based on similarity of documents

or user profiles. To improve the effectiveness of search, there is an increasing interest in personalized or user-dependent search [3]. Personalized search systems expect to deduce user search preferences received from user feedback, which is crucial in web searches and image recommendation. People often find it very challenging when searching for images as in various situations they only know which images are relevant after they see them. Their cognitive abilities can understand an image when they see it in front, but their mind has confined ability to manifest a rich object like an image. This regular conscious consumption of information leads to the problem of information overload, for information that we are interested in is much harder to locate. People, in general, reflect an image based on the images seen before. Textual and visual representations in search engine result pages (SERPs) can be perceived in the context of a seminal state within Information Foraging Theory [2, 15, 16]. Information Foraging Theory postulates that users look at those information patches which have the strongest scent, where the scent strength is estimated by textual and visual cues from the information environment, contemplating the cue's relevance to the search task. After users start interacting with text-based recommender engines they provide the system with clues about their personalized preferences. The so gathered preferences are used to increase users' visual attention to enhance personalized image recommendation. The concept of foraging intervention, which we argue can be used in explainable recommendation, refers to a task of selecting the right item from a list of recommended documents or images, presenting that these interventions can coherently shape the information scent effects of user preferences, where the correct foraging strategy not only helps inferring those preferences, but also minimising users' cognitive load.

In this paper, we have collected real image data including visual bookmarks from a popular image-based social media network (Pinterest). In order to assess user attention within the recommended images from the test collection and the effects of making a choice among them, we investigated the impact of visual bookmarks pinned to every image by determining the information scent of images. Also, we explored the foraging effects of image recommendation in terms of user engagement and satisfaction.

The contribution of this work is two-fold:

- We propose a personalized recommendation system for image search that incorporates users' visual attention to recommended items;
- We describe the user-dependent aspects we observe during foraging intervention across various effects of scent on a recommendation.

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

EARS '19, July 25, 2019, Paris, France

© 2019 Copyright held by the owner/author(s).

ACM ISBN 978-x-xxxx-xxxx-x/YY/MM.

<https://doi.org/10.1145/xxxxxx.xxxxxxx>

2 RELATED WORK

The work in this paper rests on prior research in various areas, particularly Information Foraging Theory from behavioral psychology, image-based recommender systems, and image representation and content classification from machine learning.

Information Foraging Theory: Information Foraging Theory [16] aims to model the information retrieval behavior which includes how information seekers navigate through information environments such as the web and help users finding their search strategies. Based on this theory, the user behavior to forage in the webpages (which are our information patches, see below) for specific information by trailing the information features (cues) on the Web is drawn by the patch's scent (information clues). In general, the foraging theory is based on the cost (time spent in search) and benefit (information consumption) assessments, and contemplation that people or animals recline toward rational strategies to maximize their information access or energy over an expanse of a given time. To adopt IFT for information seeking behavior which includes locating valuable pieces of information (document, image or other forms of data), seekers need to constantly evaluate cues from the online content spread over the Web. To this end, IFT follows three major concepts, which are: (i) *Information Patch* designates a physical and conceptual space [5] of information which includes a webpage or an image divided into several regions where each region¹ is made of pixels; (ii) *Information Scent* refers to the user's individual semantic compatibility to information objects and the preferred paths while navigating among/between patches via cue to estimate which nearest navigation path negotiates the probable value of distinct information object. Examples of information scent are visual or textual representations of the content i.e., text labels, tags, color or font. And (iii) *Information Diet* refers to the combined set of information that has some perceived value to a searcher, who then emulates the set of information and neglects the rest [15]. Unfavorable information is emulated if a searcher pursues a generalized diet that comprises every genre of information confronted. A searcher will then spend much time searching if the information diet is overly idiosyncratic, that is, only some genres of information are available in the information diet.

Liu et al. [13] investigated an adaptive user interaction framework by applying IFT to demonstrate the effects of image search experience based on various user types realized during the quantitative analyses of three derived evaluations. Based on different user types for content-based image retrieval, Liu et al. [12] demonstrated an IFT inspired user classification model to understand the users' interaction by functioning the model on several interaction features collected from the screen capture of various user task types on a content-based image retrieval system. They evaluated the classification model by performing qualitative data analysis and found that the six characteristics in the model are consistent with those interaction features which built a preliminary practice to study user interaction/behavior via IFT.

Personalized Image Recommendation: The recent advancements in personalized image recommendation pave the way for various image recommender systems which include image-aware and image-unaware recommendation models, specifically on social network data. These two types of image recommender systems overlay various schemes introduced in [1, 7], which efficaciously opt out images from a large collection of candidates that fit user's preference. The first type, image-aware recommender models, solely focuses on image representations and user modeling [1, 7], where representing images expressively and differentially has become one of the motives for image recommendations. However, He and McAuley [7] developed a model to exploit a pre-trained deep neural network, which supports the extraction of visual semantic embeddings from matrices containing images' pixels. This method lacks the image visual features in an immense fine-grained level because of treating an image as a whole single object. Recent work [1] proposed a multimedia recommendation model with an attention network side-by-side, which considers capturing image segments with comparative importance. In fact, this technique splits an image into equal-sized regions with the exclusion of semantic objects. We reckon that user preference to a definite image is supported by the inclusion of object semantics and exclusion of semantic objects leads to fallacy in image selection, descending the entire effectiveness of image recommendation.

On the other hand, the image-unaware recommendation models are entirely based on user modeling rather than considering the visual features of images. For instance, a user-item interaction without image information was described in [18] which introduces a pairwise learning algorithm with implicit feedback. However, few techniques are developed to pilot behavior patterns or user profiles systematically to renovate the performance of recommendations [8–10]. Past work [19] introduced a task (topic) sensitive model to characterise effects of the social network in a personalized image recommendation task. In our work we will investigate a content-based image recommendation in which we transform the image items in the representation space to recommend (or search) for identical items.

3 PERSONALIZED SEARCH RECOMMENDATION

We develop a personalized search recommendation engine for image searching based on what we call the *User-Image-Cue model*. The schematic architecture is depicted as Figure 1. In the User-Image-Cue model, users, images and (re-ranked) cues are framed on the left hand side of the figure within the interlinked graph. Their connection and role within the recommendation process is explained below. The advantage of adopting content-based recommendation over collaborative filtering is that it does not have the cold start problem [11] where a new item (or user) is introduced without previous history, as well as sufficient amount of data, whereas the latter exploits the users' correlation to make a recommendation. The content-based recommendation engine directs the image objects in the representation space, which permits to recommend for similar items.

¹generally rectangular, or could be of different shape based on the selected region of the object in an image

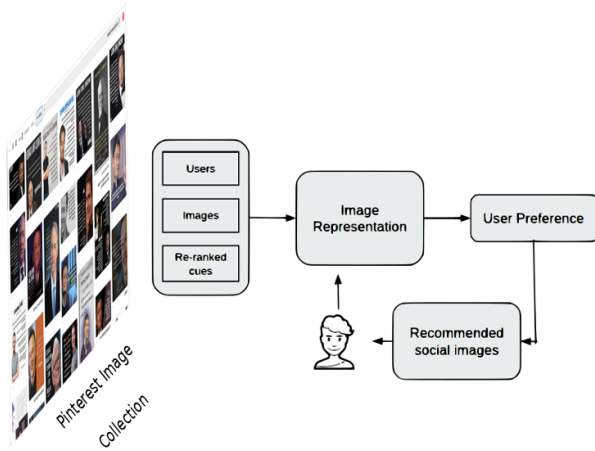


Figure 1: Personalized Image Recommendation

As per the above architecture, we make personalized search recommendations for image searching as shown in Figure 2 (which consists of four screen shots of our recommendation prototype).

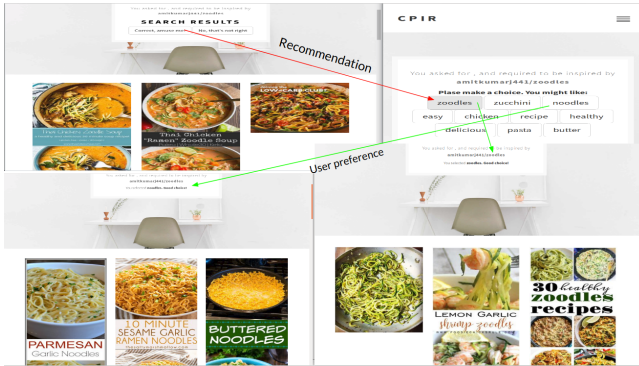


Figure 2: Personalized Search Recommendation Interface

In the first phase (top left part of the figure), we have added a Pinterest board widget in the recommendation engine where the user inputs his/her board name² as a keyword-based query and it syncs the entire image collection in real-time from the specified Pinterest board. Users seeing the search result in the second phase (top right part of the figure) will be given several preferences based on their current search result to choose from, and if chosen the recommendation system again retrieves similar items (indicated by the green arrows in the figure pointing to different results in the bottom right and bottom left, respectively). Each and every image from the collection includes a cue associated with it.

Image Representation: The image representation technique follows the hard-coded features of images which is a way to scale down computation, and as a simplified scenario of image embedding [4], similar to word embedding (word2vec) where we first

train the images with {image, label} dataset using a neural network, which then transform the image matrix representation (for instance 224x224x3) to a much smaller vector representation (image2vec). This method can be used to compute the similarity between various images (or look for close vectors that depicts similar images).

The motivation to adopt such interpretation is due to characterising the behavioral aspects of interactive elements (tags, cues and search interface, etc.) in recommendation system as opposed to focus on techniques developing a unified personalized RecSys, which is more or less based on learning or adopting features from user instead of bringing user-driven explainability in such system. Also, our proposed personalized recommendation engine can be supported with Pinterest image search algorithm [21] which has a quite better performance in terms of querying with text.

Image Features: We use different features to characterise images such as content, texture, color, and description/title. We train classifiers for each of these features on ImageNet, and employed each of them on images to extract the connected information.

We use a pre-trained ResNet50 model [6] to train our content classifier on ImageNet³ which detects almost over 1000 different objects. Also, to predict the color (classification), we apply an unsupervised k-means clustering to match predominate colors to the generic color labels using html color scheme.

4 FORAGING EFFECTS

The first work on Visual Information Foraging can be traced in [17] to find information more quickly when there is a strong information scent [2] realised from cognitive perspectives. In this paper, we apply visual Information Foraging on a personalized image recommendation scenario to understand what drives a user to a generic search result (i.e., image) in terms of user engagement and satisfaction.

We describe the effects of foraging in the context of image search which propagates implicit feedback for a content-based recommendation. In order to understand the personalized search recommendation interface by means of Information Foraging Theory, we formulate this recommendation system where the search engine result page (SERP) can be viewed as information patch together with all possible image views shapes a topology. The user aims to locate the interesting item in order to attain a decision in the foraging loop [15].

We hypothesize images as exemplary image patches⁴ that can be reached via cues while viewing an image content when it enables user cognitive beliefs. A user can activate such key beliefs via generating implicit cues to perceive ideas and plans for seeking, gathering and information consumption. In the same way animals believe on scents to forage [16] which is analogous to users following various kinds of cues in assessing image contents and navigating across patch spaces depend on images' scent.

Images and tags form cues that correspond to an information scent. To acquire more information for locating the interesting image the cues compose the information diet and information access costs. The above discussion leads to three variables that can be interpreted via the personalized image recommendation system –

²Similar to boards known in Pinterest, a board in our sense allows people to organize all their visual cues around diverse interests, ideas and plans

³<https://pjreddie.com/darknet/imagenet/#resnet50>

⁴Images patches are image regions of a particular image when treated separately

the strength of the information scent, the effort involved in making conscious consumption about the image information and the information access cost for seeking extra information about an image.

Thus, from an IFT perspective, an image I consists of n image patches $I = \{I_{p_{i,1}}, I_{p_{i,2}}, \dots, I_{p_{i,n}}\}$. For each image patch $I_{p_{i,j}}$ ($j = 1, 2, \dots, n$) we investigate those m patches whose attention by the user U is known and share the strong information scent with $I_{p_{i,j}}$. Empirically, we compute the information scent of these image patches based on the frequency of user preferences for particular content. We use a psychometric scale such as Likert scale of 1-10 with '1' being least frequently to '10' being most frequently for evaluation of the recommendation system. The information scent of every user preference based on the recommendation is reported in Table 1.

5 EXPERIMENTAL EVALUATION

5.1 Data Collection

To evaluate the proposed recommendation system (RecSys), we compiled a real image dataset from Pinterest.com, a popular visual discovery sharing platform. We collected over 1116 images belonging to two categories of foods which includes *Spaghetti Bolognese* and *Zoodles*. We split the image data into 67% train and 33% test data. The associated information labels such as title and description with the images may indicate a very complex concept, where we use Naive Bayes to count the frequency of keywords (after data cleaning process).

5.2 Results

This section reports the evaluation result of user preferences based on the personalized image recommendation. We denote information scent and recommendation by "IS" and "R" respectively. In Table 1, each recommendation (R_1, R_2, R_3, R_4, R_5) is ordered based on the strong information scent of user preferences, in which R_i represents the inferred preferences of the user based on the i -th most liked images (e.g., "Bolognese" and "Zoodles" for R_1) in the respective food categories collection. This means that those preferences with higher information scent are likely to be recommended to and attained by the searcher. This foraging-based observation makes users more likely to adopt visual bookmarks (visual cues) with little effort by hovering over recommended images instead of memorising the items themselves (with the latter discussed in [20]). This approach helps avoiding the searcher not to consume any sort of extra information diet (by memorising either items or buttons/tags).

If we interpret our scenario in terms of Information Foraging Theory, an image having either "Bolognese" or "Zoodles" (as in R_1) has a strong information scent. This means that such an image, presented as information patches, receives a comparatively large degree of attention by the user while likely to consume maximum information (information diet) and having lower information access costs, for instance in terms of the time spent on search.

6 CONCLUSION & FUTURE WORK

This paper has investigated a personalized image recommendation system from an Information Foraging perspective. To this end, we conducted an empirical evaluation of user preferences in terms of

Food Categories	Spaghetti Bolognese		Zoodles	
	User Preferences	IS	User Preferences	IS
R_1	Bolognese	10	Zoodles	9
R_2	Spaghetti	7	Zucchini	8
R_3	Recipe	6	Easy	6
R_4	Sauce	6	Pasta	5
R_5	Easy	3	Chicken	5

Table 1: Information scent of User Preferences

information scent to get some first understanding of the effects of user attention during an image recommendation scenario in the context of IFT. This work found that:

- (1) Information scent of an image has user-dependent aspects and users' scent of the same image can differ (For instance; "Bolognese" and "Spaghetti");
- (2) The overall information scent of an image (as described in [14]) becomes stronger when adding cues;
- (3) Reinforcing visual attention has a strong information scent, however, in some situations, the images' scent can exceed the cues' scent.

We intend to scale up this study on a large test collection of images with varied categories that provide human expertise for characterising the images including its applicability in an explainable recommendation. Also, this work as the preliminary practice on applying IFT in recommendation system opens the door to evaluate such scenarios on other performance measures that can shed more light on efficiency and effectiveness by exploring interactions between information scent and cue strength.

ACKNOWLEDGMENTS

This work was carried out in the context of Quantum Access and Retrieval Theory (QUARTZ) project, which has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 721321.

REFERENCES

- [1] Jingyuan Chen, Hanwang Zhang, Xiangnan He, Liqiang Nie, Wei Liu, and Tat-Seng Chua. 2017. Attentive collaborative filtering: Multimedia recommendation with item- and component-level attention. In *Proceedings of the 40th International ACM SIGIR conference on Research and Development in Information Retrieval*. ACM, 335–344.
- [2] Ed H Chi, Peter Pirolli, Kim Chen, and James Pitkow. 2001. Using information scent to model user information needs and actions and the Web. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM, 490–497.
- [3] Zhicheng Dou, Ruihua Song, and Ji-Rong Wen. 2007. A large-scale evaluation and analysis of personalized search strategies. In *Proceedings of the 16th international conference on World Wide Web*. ACM, 581–590.
- [4] Andrea Frome, Greg S Corrado, Jon Shlens, Samy Bengio, Jeff Dean, Tomas Mikolov, et al. 2013. Devise: A deep visual-semantic embedding model. In *Advances in neural information processing systems*. 2121–2129.
- [5] Peter Gärdenfors. 2004. Conceptual spaces as a framework for knowledge representation. *Mind and Matter* 2, 2 (2004), 9–27.
- [6] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. 2016. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*. 770–778.
- [7] Ruining He and Julian McAuley. 2016. VBPR: visual bayesian personalized ranking from implicit feedback. In *Thirtieth AAAI Conference on Artificial Intelligence*.
- [8] Meng Jiang, Peng Cui, Fei Wang, Wenwu Zhu, and Shiqiang Yang. 2014. Scalable recommendation with social contextual information. *IEEE Transactions on Knowledge and Data Engineering* 26, 11 (2014), 2789–2802.

- [9] Yuchen Jing, Xiuzhen Zhang, Lifang Wu, Jinqiao Wang, Zemeng Feng, and Dan Wang. 2014. Recommendation on Flickr by combining community user ratings and item importance. In *2014 IEEE International Conference on Multimedia and Expo (ICME)*. IEEE, 1–6.
- [10] Yuncheng Li, Jiebo Luo, and Tao Mei. 2014. Personalized image recommendation for web search engine users. In *2014 IEEE International Conference on Multimedia and Expo (ICME)*. IEEE, 1–6.
- [11] Blerina Lika, Kostas Kolomvatsos, and Stathes Hadjiefthymiades. 2014. Facing the cold start problem in recommender systems. *Expert Systems with Applications* 41, 4 (2014), 2065–2073.
- [12] Haiming Liu, Paul Mulholland, Dawei Song, Victoria Uren, and Stefan Rüger. 2010. Applying information foraging theory to understand user interaction with content-based image retrieval. In *Proceedings of the third symposium on Information interaction in context*. ACM, 135–144.
- [13] Haiming Liu, Paul Mulholland, Dawei Song, Victoria Uren, and Stefan Rüger. 2011. An information foraging theory based user study of an adaptive user interaction framework for content-based image retrieval. In *International Conference on Multimedia Modeling*. Springer, 241–251.
- [14] Faidon Loumakis, Simone Stumpf, and David Grayson. 2011. This image smells good: effects of image information scent in search engine results pages. In *Proceedings of the 20th ACM international conference on Information and knowledge management*. ACM, 475–484.
- [15] Peter Pirolli. 2007. *Information foraging theory: Adaptive interaction with information*. Oxford University Press.
- [16] Peter Pirolli and Stuart Card. 1999. Information foraging. *Psychological review* 106, 4 (1999), 643.
- [17] Peter Pirolli, Stuart K Card, and Mija M Van Der Wege. 2001. Visual information foraging in a focus+ context visualization. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM, 506–513.
- [18] Steffen Rendle, Christoph Freudenthaler, Zeno Gantner, and Lars Schmidt-Thieme. 2009. BPR: Bayesian personalized ranking from implicit feedback. In *Proceedings of the twenty-fifth conference on uncertainty in artificial intelligence*. AUAI Press, 452–461.
- [19] Jitao Sang and Changsheng Xu. 2012. Right buddy makes the difference: An early exploration of social relation analysis in multimedia applications. In *Proceedings of the 20th ACM international conference on Multimedia*. ACM, 19–28.
- [20] Tobias Schnabel, Paul N Bennett, Susan T Dumais, and Thorsten Joachims. 2016. Using shortlists to support decision making and improve recommender system performance. In *Proceedings of the 25th International Conference on World Wide Web*. International World Wide Web Conferences Steering Committee, 987–997.
- [21] Linhong Zhu. 2018. Demystifying Core Ranking in Pinterest Image Search. *arXiv preprint arXiv:1803.09799* (2018).