# Postprint of Internet Advertising Effectiveness Evaluation Based on Impression Space

**Authors:** Song Yongqiang, Wang Hong, Wang Lutong, Hu Xiaohong

**Date:** 2019-04-01T00:00:00+00:00

## Abstract

Evaluating the effectiveness of internet advertising is a core issue in online marketing; however, existing evaluation methods suffer from problems such as single information source, undifferentiated assumption, and global assumption, posing significant challenges to the evaluation of internet advertising effectiveness. Finding a novel evaluation metric for measuring internet advertising effectiveness has become an urgent task. First, we innovatively propose the concept of impression space as a more effective metric for evaluating web advertising effectiveness, to address the problem of single information source. Second, we analyze the influence of user characteristics such as user type, behavior patterns, and behavioral processes on internet advertising effectiveness evaluation criteria, eliminating evaluation bias caused by the undifferentiated user assumption. Third, we introduce locality features of web pages, analyzing factors such as page layout and relevance between advertisements and page content to their impact on internet advertising effectiveness, thereby eliminating the global assumption. Finally, we construct an impression space model based on multimodal features to predict internet advertising effectiveness. Experimental results demonstrate that the proposed impression space significantly improves the accuracy of internet advertising quality evaluation, reaching 92.4%. Moreover, the prediction results of the impression space model are not only more accurate and scientific, but also exhibit significant interpretability.

## Full Text

## Preamble

**Title:** Evaluation of Internet Advertising Effectiveness Based on Impression Space

**Authors:** Song Yongqiang[1,2], Wang Hong[1,2]†, Wang Lutong[1], Hu Xiaohong[1]
[1](College of Information Science & Engineering, Shandong Normal University,

Jinan 250358, China; [2]Shandong Provincial Key Laboratory for Distributed Computer Software Novel Technology, Jinan 250014, China)

**Abstract:** Internet advertising effectiveness evaluation is the core issue of online marketing. However, current evaluation methods suffer from problems such as single information sources, undifferentiated assumptions, and global assumptions, posing significant challenges to accurate assessment. Finding a new evaluation metric for Internet advertising effectiveness has become an urgent task. This paper first innovatively proposes the concept of impression space as a more effective metric for webpage advertising effectiveness evaluation to address the single information source problem. Second, it analyzes how user types, behavioral patterns, and behavioral processes affect evaluation criteria, eliminating biases caused by undifferentiated user assumptions. Third, it introduces local webpage features to analyze how page layout and advertisement-content relevance influence advertising effectiveness, thereby eliminating global assumptions. Finally, it constructs a multimodal feature-based impression space model to predict Internet advertising effectiveness. Experimental results demonstrate that the proposed impression space significantly improves evaluation accuracy to 92.4%. Moreover, the impression space model' s predictions are not only more accurate and scientific but also exhibit clear interpretability.

**Keywords:** impression space; user behavior; area of interest; Internet advertising effectiveness

## 0 Introduction

Internet advertising, as a product of the Internet era, represents marketing information disseminated through network services. With its strong interactivity, broad reach, and flexible timeliness, it has become an essential component of modern advertising media. Recent research indicates that Internet advertising in the new media technology era has permeated all aspects of contemporary social life. The goals of Internet advertising are to improve advertisers' return on investment, enhance user experience, and achieve multi-win outcomes across the industry chain. The key to achieving these objectives lies in more precise ad placement, which fundamentally depends on accurately evaluating advertising effectiveness. Current evaluation methods face significant challenges due to three main problems: single information sources, undifferentiated assumptions, and global assumptions.

First, existing Internet advertising effectiveness metrics are simplistic, primarily relying on click-through rate (CTR) and conversion rate. CTR's major drawback is its failure to capture ads that users may notice but do not click. Conversion rate, meanwhile, conflates webpage browsing with ad browsing.

Second, current evaluations ignore user differences, considering only active user behaviors. In reality, variations in user types, thinking patterns, and browsing habits substantially impact advertising effectiveness.

Third, evaluations operate under a global assumption, assessing advertising effectiveness across entire webpages without considering fine-grained differences such as ad placement and advertisement-content relevance.

This paper argues that webpage advertising effectiveness is determined by multiple factors beyond the ad itself, including user type, psychological characteristics, page layout, and browsing trajectories. Therefore, we propose impression space as an evaluation standard for online advertising effectiveness and implement a corresponding evaluation method. The main contributions are:

a) To address the single information source problem, we innovatively propose the concept of impression space as a more effective webpage advertising effectiveness metric.

b) To counter undifferentiated user assumptions, we analyze how user types, behavioral patterns, and behavioral processes affect evaluation criteria, eliminating resulting biases.

c) To eliminate global assumptions, we introduce local webpage features to analyze how page layout and advertisement-content relevance influence advertising effectiveness.

d) Based on the above, we construct a multimodal feature-based impression space model to predict Internet advertising effectiveness, yielding more accurate, scientific, and interpretable results.

---

## 1 Related Work

Compared to traditional mass media advertising, Internet advertising offers broader coverage, stronger interactivity, and lower costs, playing a pivotal role in online marketing systems. Consequently, Internet advertising research has attracted increasing attention.

Currently, click-through rate serves as the primary evaluation metric for Internet advertising effectiveness. Notable click models include the Location-based Click Model (LCM) [?], Vertical-aware Click Model (VCM) [?], Temporal Click Model (TCM) [?], and mathematically-based click models. LCM verifies that users' browsing sequence aligns with search result positions. VCM demonstrates that different vertical search results significantly affect user gaze behavior, accurately reflecting user attention patterns. TCM describes users' non-sequential clicking behavior, indicating that browsing may be non-ordered. Among mathematically-based models, Shan et al. [?] utilized factorization of fully coupled interaction tensors to predict CTR. Wang et al. [?] employed a Partially Observable Markov (POM) model to predict click sequences, mathematically proving the model' s accuracy to some extent. However, these studies focus solely on active user behaviors, neglecting unconscious behaviors during browsing, resulting in incomplete feedback on users' true intentions and biased advertising effectiveness evaluations.

With the maturation of deep learning architectures, Qu et al. [?] proposed a Product-based Neural Network (PNN) model by introducing a product layer between embedding and fully connected layers. To better address feature combination issues, Shan et al. [?] proposed a Deep Crossing model that automatically combines features. These methods can be summarized as using deep network architectures to automatically learn data features and enhance data representation capabilities. Although deep networks can effectively utilize structured neural network models to adapt to complex nonlinear relationships in advertising datasets, they ignore users' actual feedback in click prediction, compelling researchers to seek metrics beyond CTR for more accurate Internet advertising effectiveness evaluation.

Consequently, a series of advanced Internet advertising effectiveness evaluation methods have emerged. Zhao et al. [?] applied the entropy-weighted double-base point method to evaluate Internet advertising effectiveness, effectively solving information bias problems. Deng et al. [?] combined reach rate with CTR to evaluate effectiveness through effective reach rate representing audience ratios. Lyu et al. [?] studied Internet advertising effectiveness through user browsing patterns, while Chen et al. [?] analyzed user-ad interaction behaviors, summarizing advertising effects associated with popular terms. Despite their respective focuses, these studies struggle to define a comprehensive metric for actual advertising effectiveness. Subsequently, Hu et al. [?] proposed using memorability and interest to measure effectiveness, but ignored differences in how users receive, store, transform, extract, and utilize advertisements. Wedel et al. [?] noted that individual differences create varying impressions of products. Dreze et al. [?] also pointed out CTR' s limitations, arguing that users' ad browsing paths vary by personal experience. Unfortunately, Dreze' s work relied on traditional experiments lacking in-depth analysis with professional equipment.

The second type of difference concerns how webpage layout characteristics affect user information processing feedback. Wang et al. [?] obtained different users' common browsing patterns through a Partially Sequential Click Model (PSCM), but failed to consider feedback differences from ads in different webpage regions and inter-user variations.

In user browsing behavior research, most researchers use eye-tracking data as a baseline, considering data close to human eye browsing as ground truth. Eye-tracking data can identify user attention points and browsing trajectories, authentically reflecting areas and activities of interest and exploring unconscious behaviors. However, existing literature focuses primarily on natural image processing, with limited webpage research. Renshaw et al. [?] studied graphic design through eye-tracking experiments, while Shen et al. [?] investigated webpage ad saliency, achieving only 0.7206 AUC. With advancing eye-tracking technology, experiments can now 摆脱 heavy mechanical equipment, enabling research beyond theoretical aspects. Xu et al. [?] applied eye-tracking data to video processing, while references [?, ?] fused eye-tracking and mouse data to predict user satisfaction with search results. Therefore, this paper employs eye-tracking experiments to study user gaze trajectories, analyze implicit information such as psychological activities, discover browsing patterns among different user types, obtain implicit behaviors, identify user interest areas, propose more objective and accurate online advertising effectiveness evaluation metrics, and provide foundations for precise Internet ad placement.

---

## 2 Impression Space

We propose impression space as an evaluation metric for Internet advertising effectiveness. Social psychologists indicate that impression formation is a process where subjects judge object characteristics, influenced by both cognitive subjects and object feedback information (expressed via word vectors as impression space = (subject information, object feedback information)).

Psychological research shows impression formation may be affected by cognitive style, emotion, and object feedback information. Cognitive styles can be divided into independent and dependent; user emotional states into mood and passion; object feedback information into complete feedback, possible feedback, and failed feedback. We define impression space as expressed by 2 (styles) $\times$ 2 (emotional states) $\times$ 3 (object feedback information), as shown in Table 1 .

---

## 3 User Type Classification

To eliminate undifferentiated assumptions, this paper classifies user types and analyzes their impact on Internet advertising effectiveness evaluation.

### 3.1 User Combined Attributes

Attribute combinations are diverse (e.g., personality, behavioral habits). We utilize common distributional combinations from mathematical statistics. Taking normal distribution as an example, we explore attribute column distributions.

Definition 2: Attribute Weight. For a sample set D divisible into k clusters, where the l-th cluster is represented as, the single attribute weight is, where is the weight of the t-th attribute in the l-th cluster. The combined attribute weight is the weight of the j-th combined attribute in the l-th cluster.

If any attribute set follows the same distribution and satisfies the given distribution function F, defined as, for single attribute weight and combined attribute weight respectively satisfying:, given constraints, satisfying the central limit theorem, where and are the overall attribute mean and variance. The central limit theorem indicates that when attribute data meets certain conditions, random variables approximately follow normal distribution. If attribute satisfies the central limit theorem, its distribution can be described by, i.e., the probability density function of attribute.

Attributes satisfying normal distribution probability density functions can be classified into combined attributes of normal distribution. Other probability density functions are shown in Table 2 . If any attribute column satisfies a probability density function distribution (unknown quantities may differ but distribution patterns are consistent), it is classified as a combined attribute.

### 3.2 User Type Classification Based on Attribute Combination

User type classification follows two approaches. The first is directly determined by attributes, as shown in Figure 1 [Figure 1: see original paper]. The second is jointly determined by single and combined attributes, where the first-layer contribution is called single attribute contribution and the second-layer contribution is called combined attribute weight, as shown in Figure 2 [Figure 2: see original paper].

---

## 4 User Behavior

We investigate how different user behavioral patterns affect Internet advertising effectiveness using frequent sequence algorithms and path merging algorithms.

### 4.1 Click Behavior

We employ the Apriori algorithm to discover user click behavior association rules. Based on Apriori' s prior law (if a set is frequent, all its subsets are frequent), we merge partial mining results to facilitate observation of click pattern association rules and study user click preferences. Fundamental concepts in association rules are support and confidence.

Definition 3: Support. The frequency of a frequent itemset or rule occurring across all transactions determines how frequently the rule applies to the given dataset. Support typically eliminates meaningless rules. Association rule support is defined as.

Definition 4: Confidence. The frequency of event Y occurring in transactions containing event X measures inference reliability through rules. Association rule A B confidence is.

Definition 5: Frequent Itemset. In association rule mining, sets satisfying certain minimum confidence and support thresholds are called frequent sets.

### 4.2 Browsing Behavior

We identify user browsing preferences through path merging algorithms based on browsing pattern decomposition (e.g., any user browsing pattern path 1→2→3 can be decomposed into subpaths 1→2 and 2→3). This approach mines user subpaths and merges them to obtain user browsing patterns.

Definition 6: User Browsing Matrix S. Based on the search engine webpage layout format (10 URLs for main content plus ad area and right-side area), we establish a 12×12 matrix where element values represent page visit counts (e.g., a value of 4 indicates subpath 1→2 appears 4 times across all browsing patterns).

Definition 7: Multi-frequency Subpath. Given user browsing matrix S, we calculate row sums as, where. Processing matrix S yields matrix, where element values are. Defining a threshold S, if matrix element value.

Definition 8: Path Merging. If m→n and n→l are multi-frequency subpaths, they merge into m→n→l, where m, n, and l can be single page layout positions or other multi-frequency subpaths.

---

## 5 Interest Area

We propose dividing interest areas based on fixation duration and page layout to eliminate global assumption impacts on Internet advertising effectiveness.

### 5.1 Page Layout Characteristics

Page layout typically consists of 10 URLs (main content) plus advertisement and right-side areas, as shown in Figure 3 [Figure 3: see original paper] (L1, L2, and L3 represent ads appearing above, below, and to the right of main content, collectively called the AD area). Users develop inherent browsing habits due to visual characteristics and personal browsing patterns. In practice, obtaining more user browsing information is difficult. Researching entire webpage information is not only cumbersome but also time-consuming, motivating researchers to identify interest areas with sufficient valuable information and sufficiently small scope.

### 5.2 Interest Area Based on Fixation Duration

Definition 9: For any area i, fixation time is, where Ftime is fixation duration, Dtime is dwell time, and Fcount is fixation count. The fixation duration formula for any area i is.

Using fixation duration and page layout to divide interest areas, we retain advertisement area AD and right-side area R as two separate research regions. Based on fixation duration calculations, we use the average fixation duration across 10

URL areas as a baseline. URL areas with fixation duration above the baseline are classified as interest areas, while those below are non-interest areas.

---

## 6 Data Collection

This paper proposes the impression space concept, but currently lacks datasets linking eye-tracking data with user impression spaces. Therefore, we first conduct large-scale data collection.

### 6.1 Experimental Subjects

We randomly recruited undergraduate students (aged 18-21) from various majors at our university as volunteers. All participants had no eye diseases or related conditions, meeting eye-tracking experiment requirements.

### 6.2 Experimental Equipment

The eye-tracking experiment used eye-tracking equipment Version 2.4 developed by German SMI Technology, with a sampling frequency of 120 Hz. Recording and analysis employed built-in software IViewX, Experiment Center, and BeGaze. Data processing and analysis used Spyder and Java.

### 6.3 Experimental Materials

This study required collecting both user subject information and user browsing information.

User subject information was collected using the Embedded Figures Test (EFT) for cognitive style classification and a psychological state test table proposed by renowned American psychologist Witk et al. for emotional state classification.

User browsing information used webpage materials. We selected diverse product types (luxury goods, tickets, home appliances, etc.) covering broad categories to simulate users' daily selection mindsets, with two different brands for each type. We chose Baidu as the search engine because it is the world' s largest Chinese search engine with layout similar to most Chinese search engines. To ensure realistic experimental effects, participants browsed freely as usual, including clicking and scrolling. To control experimental variables and ensure participants saw identical SERPs for each product type, we crawled pages from the search engine and retained only commercial promotions in research positions. After browsing, we tested participants' memory to describe object information feedback. Eye-tracking equipment recorded participants' eye movement information, while embedded JavaScript code captured click information.

### 6.4 Experimental Procedure

a) Collect user subject information. The test required participants to find and trace a specified simple figure hidden within complex graphics, consisting of 25 questions with increasing difficulty.

b) Collect user browsing information. Before webpage browsing, standard eye-tracker calibration was performed. Formal experiments began after participants met calibration standards. Participants were informed of product types via on-screen prompts (e.g., "mobile phones" ), adopting a buyer' s mindset for that product type. The browsing process mirrored daily web browsing—participants could click items of interest. To prevent fatigue, each participant viewed a maximum of 6 random SERPs and could stop the experiment at any time. Post-experiment memory data was collected.

c) Impression space information. Impression space was determined through user style test questions, browsing emotional states, and webpage feedback information.

---

## 7 Impression Space Analysis

CTR has long been the gold standard for Internet advertising evaluation. However, marketers' mass ad deliveries for higher economic returns have degraded advertising quality, reducing user interest. With average webpage CTR below 1%, CTR struggles to reflect true advertising effectiveness. Latest data shows average online ad CTR has dropped from 30% to below 0.5%. For brand advertising, success depends less on clicks or post-reading purchases and more on whether users form lasting impressions, create brand effects, establish unique positive brand/product images, and improve long-term offline conversion rates.

### 7.1 Impression Space Model Features

Using impression space as class labels, we treat experimentally collected data as model features. Table 3 shows features affecting user impression space formation.

Regarding class labels for impression space, the quantification method is as follows:

Definition 10: Cognitive style = independent (1), dependent (0); Emotion = mood (1), stress (0); Feedback information = complete feedback (2), partial feedback (1), failed feedback (0). The vector expression for impression space labels is:

- Existing impression: [(1,1,0); (0,1,2); (1,0,2); (1,0,1); (0,0,2)]
- Fuzzy impression: [(1,1,1); (0,0,1)]

- Missing impression: [(1,1,0); (1,0,0); (0,1,1); (0,1,0); (0,0,0)]

### 7.2 Effectiveness Validation

To further validate the proposed impression space' s effectiveness, we conducted data collection and annotation. The experiment prepared webpage materials via web crawling, collected multi-user browsing and ad click information, and used chi-square testing for validation. Chi-square testing is a common hypothesis testing method for detecting deviation between predicted and actual values. Larger chi-square values indicate greater deviation and poorer prediction methods, while smaller values indicate better methods. This study treats expert scale annotations as actual values and impression space retention ability and click status as predicted values.

Using impression space and expert scale annotations, chi-square testing proceeds as:

a) Propose null hypothesis H : No deviation exists between impression space ability and expert scale annotations, with Sig = 0.05.

b) Calculate theoretical values.

c) Calculate $^2$ value.

d) Consult $^2$ table.

At Sig = 0.05, we cannot reject H , indicating both methods are consistent.

### 7.3 Impression Space Model Construction

For the impression space model, we propose two structural stacking methods. The first method, based on user types and preferences, is shown in Figure 4 [Figure 4: see original paper].

The second method is shown in Figure 5 [Figure 5: see original paper].

### 7.4 Experimental Results

To demonstrate the proposed method' s effectiveness, we conducted three experiment groups: (1) selecting optimal user type numbers to maximize inter-group differences and analyze user behavior; (2) constructing optimal impression space models; (3) comparative result demonstration.

**7.4.1 User Behavior Analysis** Using Robert' s within-group sum of squares method to determine optimal user type quantity, results are shown in Figure 6 [Figure 6: see original paper], with clustering results in Figure 7 [Figure 7: see original paper]. Results indicate two user types as the optimal cluster number, showing significant inter-type differences that facilitate mining behavioral dynamic information like browsing patterns and attention points.

Using frequent sequence mining for click preference association rules under different user types, results are shown in Table 4 . Results show both user types click top-down, mostly sequentially but with some return visits. Top-down clicks indicate initial interest, prompting subsequent clicks; when interested, users may click back-and-forth for comprehensive understanding.

Path mining algorithms yielded different user type browsing patterns, as shown in Table 5 . Results indicate browsing sequences generally satisfy the vertical assumption of top-down scanning.

Using fixation duration calculations, user fixation timetables across different areas are shown in Table 6 . Combining fixation duration with page layout yields user interest areas, as shown in Figure 8 [Figure 8: see original paper].

User browsing behavior characteristics include: (a) attention hotspots form an F-pattern, decreasing from top to bottom and left to right; (b) fixation clusters concentrate in the first two-thirds of pages, typically the earliest displayed content; (c) overall left-side attention time exceeds right-side. However, this analysis is abstract and requires further behavioral analysis.

**7.4.2 Impression Space Model** Using classic machine learning classifiers as base models combined with decision tree branching screening to construct impression space models, results are shown in Figure 9 [Figure 9: see original paper]. Multi-model construction increases workload, but random forest-based impression space models already achieve near-optimal accuracy. While multi-model construction can slightly improve accuracy, it substantially increases work time and resource consumption. Therefore, we select decision tree branching combined with random forest as the base model.

Further optimization verifies whether different classification base models through multi-model architecture can produce better results, as shown in Figure 10 [Figure 10: see original paper].

**7.4.3 Model Comparison** Current Internet advertising effectiveness metrics primarily use CTR and conversion rate. CTR-based models include Partially Observable Markov (POM) model, Polynomial2 (PLOY2) model, Factorization Machine (FM) model, and Field Factorization Machine (FFM) model. Using identical datasets, we compare impression space models against click models, with accuracy and stability results shown in Table 7 .

Analysis reveals: (a) increasing user sophistication makes click behavior less reflective of true intent, with unconscious clicks affecting results; (b) many users may have positive ad impressions without taking action. In summary, CTR struggles to reflect true advertising effectiveness, motivating researchers to seek alternative metrics.

Experimental results show random forest-based impression space models perform optimally. Random forest, as an ensemble algorithm, combines base classifiers into a strong classifier through node splitting feature selection. Its randomness in sampling and feature selection, combined with precise screening strategies, achieves optimal results.

---

## References

[1] Craswell N, Zoeter O, Taylor M, et al. An experimental comparison of click position-bias models [C]//Proc of the 1st International Conference on Web Search and Data Mining. New York: ACM Press, 2008: 87-94.

[2] Wang Chao, Liu Yiqun, Zhang Meng, et al. Incorporating vertical results into search click models [C]//Proc of the 36th International ACM SIGIR Conference on Research and Development in Information Retrieval. New York: ACM Press, 2013: 503-512.

[3] Xu Wanghong, Manavoglu E, Gantu-Paze E. Temporal click model for sponsored search [C]//Proc of the 3rd ACM International Conference on Web Search and Data Mining. New York: ACM Press, 2010: 211-220.

[4] Shan Lili, Lin Lei, Sun Chengjie, et al. Predicting ad click-through rates via feature based fully coupled interaction tensor factorization [J]. Electronic Commerce Research Applications, 2016, 16 (C): 30-42.

[5] Wang Chao, Liu Yiqun, Wang Meng, et al. Inferring search behaviors using partially observable Markov model [C]//Proc of the 3rd ACM International Conference on Web Search and Data Mining. New York: ACM Press, 2010: 211-220.

[6] Qu Yanru, Cai Han, Ren Kan, et al. Product-based neural networks for user response prediction [C]//Proc of IEEE International Conference on Data Mining. Piscataway, NJ: IEEE Press, 2017: 1149-1154.

[7] Shan Ying, Hoens T. Ryan, Jiao Jian, et al. Deep crossing: Web-scale modeling without manually crafted combinatorial features [C]//Proc of the 22nd ACM SIGGKDD International Conference on Knowledge Discovery and Data Ming. New York: ACM Press, 2016: 255-262.

[8] Zhao Huifang, Zhang Yan. Entropy-weighted double-base point method in network advertising effectiveness evaluation [J]. Statistics and Information Form, 2006, 21(4): 84-86.

[9] Deng Wenfeng, Zhou Chaomin. On measurements of online advertisement effectiveness [J]. Shanghai Management Science, 2005, 27(4): 13-15.

[10] Lyu Hongming, Chen Ming. A study of online advertising effect appraisement from a custom psychology perspective [J]. Journal of Nanjing Agricultural University (Social Sciences Edition), 2006, 6(4): 46-50.

[11] Chen Lei, Liu Yiqun, Ru Liyun, et al. Performance evaluation of online sponsored search based on user log analysis [J]. Journal of Chinese Information Processing, 2008, 22(6): 92-97.

[12] Hu Xiaohong, Wang Hong, Ren Yanju, et al. Research on network advertisement effect based on eye movement technology [J]. Application Research of Computers, 2018, 35(5): 1345-1349.

[13] Wedel M, Pieters R. Eye fixations on advertisements and memory for brands: a model and findings[J]. Marketing Science, 2000, 19(4): 297-318.

[14] Dreze X, Hussherr F X. Internet advertising: Is anybody watching? [J]. Journal of Interactive Marketing, 2003, 17(4): 8-23.

[15] Wang C, Liu Y Q, Wang M, et al. Incorporating nonsequential behavior into click models [C]//Proc of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval. New York: ACM Press, 2015: 211-220.

[16] Renshaw J A, Finlay J E, Tyfa D, et al. Understanding visual influence in graph design through temporal and spatial eye movement characteristics [J]. Interacting with Computers, 2004, 16(3): 557-578.

[17] Shen Chengyao, Huang Xun, Zhao Qi. Predicting eye fixations on webpage with an ensemble of early features and high level representations from deep network [J]. IEEE Trans on Multimedia, 2015, 17 (1): 2084-2093.

[18] Xu Ying, Xu Suting, Liu Xingtong, et al. Analysis on the validity of eye movement data during video viewing [C]//Proc of International Congress on image and Signal Processing, Biomedical Engineering and Informatics. 2017.

[19] Chen Ye, Liu Yiqun, Zhou Ke, et al. Does vertical bring more satisfaction: predicting search satisfaction in a heterogeneous environment [C]//Proc of the 24th ACM International Conference on Information and Knowledge Management. New York: ACM Press, 2015: 1939-1942.

[20] Liu Yiqun, Chen Ye, Tang Jinhui, et al. Different Users, Different Opinions: Predicting Search Satisfaction with Mouse Movement Information [C]//Proc of the 38th ACM SIGIR International Conference on Research and Development in Information Retrieval. New York: ACM Press, 2015: 493-502.

[21] Chen Xiaojun, Ye Yunming, Xu Xiaofei, et al. A feature group weighting method for subspace clustering of high dimensional data [J]. Pattern Recognition, 2014, 45(1): 434-446.

[22] Clifton C, Ferreira F, Henderson J M, et al. Eye movements in reading and information processing: Keith Rayners 40 year legacy [J]. Journal of Memory and Language, 2016, 86: 1-19.

[23] Tibshirani R, Walther G, Hastie T, et al. Estimating the number of clusters in a data set via the graph statistic [J]. Journal of the Royal Statistical Society, 2001, 63(2): 411-423.

[24] Deng Xinyang, Liu Qi, Deng Yong, et al. An improved method to construct basic probability assignment based on the confusion matrix for classification problem [J]. Information Sciences, 2016, 340: 320-331.

[25] Chen Chen, Hou Chunyan, Xiao Jiakun, et al. Purchase Behavior Prediction in E-Commerce with Factorization Machines [J]. IEICE Trans on Information and Systems, 2016, 99 (1): 270-274.

*Note: Figure translations are in progress. See original paper for figures.*