Intelligent Web Advertisement Based on Eye-Tracking and Machine Learning

Yasaman Mashhadi Hashem Marandi
School of Mathematics, Statistics and Computer Science,
College of Science, University of Tehran, Tehran, Iran
marandi.yasaman@ut.ac.ir

Hedieh Sajedi*
School of Mathematics, Statistics and Computer Science,
College of Science, University of Tehran, Tehran, Iran
hhsajedi@ut.ac.ir

Received: 2021/01/28 Revised: 2021/04/10 Accepted: 2021/05/22

***Abstract*— Building and maintaining brand loyalty is a vital issue for market research departments. Various means, including online advertising, helps with promoting loyalty to the brand amongst users. The present paper studies intelligent web advertisements with an eye-tracking technique that calculates users' eye movements, gaze points, and heat maps. This paper examines different features of an online ad and their combinations, such as underlining words and personalization by eye-tracking. These characteristics include underlining, changing color, number of words, personalizing, inserting a related photograph, and changing the size and location of the advertisement on a website. They help advertisers to improve their ability to manage the ads by increasing users' attention. Moreover, the current research argues the impact of gender on users' visual behavior for advertising features in different Cognitive Demand (CD) levels of tasks while avoiding interruption of users' cognitive processes with eye-tracking techniques. Also, it provides users the most relevant advertisement compatible with CD level of a task by Support Vector Machine (SVM) algorithm with high accuracy. This paper consists of two experiments that one of them has two phases. In the first and second experiments, a news website alongside an advertisement and an advertising website is shown to the users. The results of the first experiment revealed that personalizing and underlining the words of the ad grabs more attention from users in a low CD task. Furthermore, darkening the background promotes users' frequency of attention in a high CD task. By analyzing the impact of gender on users' visual behavior, males are attracted to the advertisement with red-colored words sooner than females during the high CD task. Females pay more prolonged and more frequent attention to the ads with red-colored words and larger sizes in the low CD task. The second experiment shows that the gazing start point of users with a right to left mother tongue language direction is mainly in the middle of the advertising website.**

Keywords*— Web Advertisement, Eye-Tracking, Intelligent Advertisement, Online Advertisement, Human-Computer Interaction, Machine Learning, Website Design, User Experience.

1. INTRODUCTION

Nowadays, although people are bombarded with advertisements (ads), their attention to online advertising is lower than other media like TV and magazines [1]. At the same time, advertisers can succeed if they attract large communities. The Internet is a highly competitive environment for them, and studying advertising features has a crucial role. These features include size, color, number of words, underlining, etc. Among these features, personalization is an impactful attentional enhancing attribute. It consists of inserting individuals' information like name, address, and job in the announcement [2]. It should be considered that a possible sense of privacy invasion from users might cause adverse effects on them [2]. Therefore, studying the amount of personalizing and its impact on users is beneficial.

In recent years, social media are the most common platforms of advertising [3]. Market research departments take several approaches, including website and social media advertising, to attain higher brand loyalty. For instance, considering prominent banks' Facebook pages, the repertory grid method obtains users' preferences to implement a superior design for them [4]. The results contribute insights to social media consultants in managing the contents of a Facebook branding page. Thus, improvement in social media advertising leads to higher engagements with a massive worldwide population.

Moreover, designers are aware that the design of a website itself affects users' visual attention. Therefore, creating a website is essential for advertisers to detect users' visual behavior and adjust their ads [5]. For example, the System Usability Scale (SUS) and the eye-tracking technique can assess the performance of public sector websites on usability metrics [6]. Furthermore, they declare that minimal designs have a better user experience than the traditional ones and parts of the most alluring website.

This research studies users' reactions to various ads in two tasks with different CD levels, designing an advertising website while applying the eye-tracking technique. The studied advertising features of this paper are chosen based on the conventional banner design methods; however, personalization and combination of these features are new aspects of ads proposed by this research. CD-level of a task plays an important role in users' visual behavior toward an ad [3]. Therefore, it is essential to consider this element. Since the eye-tracking technique [2] prevents interruption of users' cognitive processes, it is used to detect users' interactions with stimuli and provides visual behavior metrics. Moreover, a machine-learning algorithm is used to classify the CD level of tasks based on users gazing pattern datasets. For these purposes, the questions of this paper are:

1) How do nine features of an ad and their combinations help attract users' attention sooner, longer, and more frequently in two levels of CD tasks?
2) How do CD levels of tasks affect users' gaze patterns?

3) How does the right to left direction mother tongue language affect interaction with an advertising website?

To analyze the experimental results, statistical methods including one-way and two-way Analyses Of Variance (ANOVA), t-tests [7-9], and SVM algorithm are applied to the data.

Contributions of this paper are:

- Interpreting users' visual behavior while viewing an ad with various features alongside an online news article in high and low CD tasks
- Recognizing the impact of gender on users' visual behavior in low and high CD tasks
- Classifying CD level of tasks according to users' gaze pattern to adjust features of an ad incompatible with their level of CD tasks
- Interpreting gaze pattern for users' with the right to left direction mother tongue language while viewing an English advertising website

This paper is presented as follows: First, it briefly reviews features of ad, types of tasks, design of websites, usability importance, and eye-tracking technique as developing hypotheses. Second, it explains measures and methods. Finally, it demonstrates the results of the experiments and discusses the findings.

2. LITERATURE REVIEW

This research analyzes the features of online ads to promote visual attention paid to them. Also, this paper studies the design of an advertising website to improve the user experience. The previous researches on advertising and non-advertising topics concerning this goal are summarized in Table I. This table presents research works categorized by their publishing years and their detailed results to this study.

2-1. FEATURES OF ADS

Ads have various attributes such as size, location, picture, color, number of words, and emotional contents [18]. There is a lack of research on these features, and this paper studies some of the noted characteristics. In studies about advertising methods, personalization is outstanding in attracting users. An ad is personalized by inserting individual information like identification information (e.g., name and email) and personal information (e.g., visited websites and hobbies) [19].

Since personalized advertising provides perceived relevance to the self [20,21], it attracts users [22,23] and promotes their attitudes toward it [24,25]. Although this feature is not novel anymore, users' information recordings can provide a more improved personalized ad. Moreover, a consumers' perceived personalization contributes to a better prediction for positive attitude effects [25-27]. A deliberate change in a generic message based on the accumulated data of the users and delivering it to them is, in theory, an actual personalization.

Furthermore, perceived personalization is defined as seeing a preferred message or a non-preferred one. Hence, the recipient has control over the perceived personalization, whereas the sender is responsible for the actual personalization [28]. Inserting users' first name in the ad increases the purchase intention of users for the recommended product of it [27]; however, in some cases, it can undo the positive impression [29-31]. Researchers have realized that it causes avoidance due to invading privacy for some users [26].

In general, advertising attitude is higher for females than males [32,33]. People pay considerably more attention to media advertising than internet advertising [1]. Hence, online advertising is in great competition with other media, so the features that affect visual attention are valuable for advertisers [2]. Not many researchers have studied the amount of personalization attentional effect on users until today. One of them concluded that inserting a photograph of people in the ad would double the time spent on looking at the ad than putting their name on it [22]. The other studies the attentional difference of people for a personalized ad compared with a non-personalized one in two different tasks [2]. The current study compares many attentional grabbing features of attentional importance, such as personalizing of an ad.

2-2. TYPES OF TASKS IN THE ASPECT OF CD

A key factor while studying visual attention paid to an ad is the users' task CD level. Based on cognitive load theory [34,35], there are limited cognitive resources at a time to process information. Therefore, while users perform a main high-level CD task (e.g., finding information), not many resources are left available for other tasks (e.g., ads). In consequence, consumers' cognitive capacity influences ad acceptance [2]. Another factor is users' Internet motive since it results in users' cognitive effort for a task [36]. For instance, users with entertainment goals are less focused on their tasks than the ones with searching goals [37,38]. Therefore, adjusting the features of an ad with each type of task is a practical approach.

The first phase of this study examines four features of users' visual behavior by eye-tracking technique. These features include 1) fewer words, 2) underlining words, 3) coloring words with a warm color (e.g., red), and 4) personalizing. In the second phase, the first phase is developed in terms of combining features and enhancing them. This phase studies tasks with two levels of CD (low vs. high) and nine different features of an ad. The features are 1) fewer words, 2) underlining words, 3) coloring words with a warm color (e.g., red), 4) personalizing, 5) personalizing and underlining words, 6) containing a photo of a person related to the advertised product (e.g., tea), 7) darkening the background, 8) changing the location of the ad (e.g., placing it on the right side of the webpage) and 9) larger size of the ad. Although these features might seem simple, the eye-tracking technique reveals that they have various effects on users' visual behavior depending on the tasks' CD levels. These effects are differentiations in amounts of attentional grabbing, at first sight, frequency, and the total length of the fixations. This paper examines each of the previous features in high and low CD levels and participants' gender effects on their visual behavior. Table II displays the hypotheses of the second phase, which contends these differentiations and variations in results depending on users' gender.

TABLE I. COMPARISON BETWEEN CURRENT RESEARCH AND PREVIOUS RESEARCHES ON ADVERTISING AND NON-ADVERTISING TOPICS CONCERNING ANALYSES OF ADVERTISING FEATURES TO PROMOTE VISUAL ATTENTION PAID TO THEM

Year	*Results of researches*
2016 [3]	•Personalized ads attract users significantly longer and more frequently than non-personalized ads. •When people are engaged in a high CD task, they tend to pay much deeper and longer attention to personalized ads than non-personalized ones. •When people are involved in a low CD task, they pay a relatively low amount of attention to personalized and non-personalized ads. •There is no significant interaction between personalization and CD of the task on perceived goal impediment and attitude toward the ad.
2010 [10]	Studying pictures and content features of the ads reveal that: •Images can make online ads appear more attractive to consumers. •As consumers easily understand the ads without involving large amounts of texts, consumers are more attracted. •Short and concise messages contribute to the effectiveness of banner ads.
2010 [11]	•Many of the theories and models developed for advertising effectiveness, and the variables used to measure it, could and should be applied when assessing the quality of the user experience while using websites in general, regardless of whether they contain ads.
2004 [12,13]	Five types of ad formats, including banners, interstitials and pop-ups, sponsorships, hypertext links, and websites, are found on the Internet.
2000 [14]	•When an individual log onto the Internet intending to shop, those items on the shopping list will do well with attention, attitude, and intent to purchase. Other things will be less effective, regardless of the structures of the interactive ads. •When consumers are motivated to use the Internet to surf or entertain themselves, it is more likely that advertising features will be better predictors of whether attention and selection of that advertising occur.

In summary, users' visual behavior relation with each type of task and ad is recognized in phase two. Therefore, presenting the best ad to the users after intelligently realizing their kind of task is essential. To classify the tasks, a dataset of users' first five gaze points is made. These points are labeled as their location on the webpage, which is divided into nine equal areas. Regarding the size and sequential format of this dataset, the SVM algorithm is used.

Moreover, to see whether there is any relation between mother tongue language direction and visual pattern, this research studies user with the right to left direction mother tongue language interaction with advertising websites by eye-tracking.

2-3. DESIGN OF WEBSITES

While reading, users pay more attention to the left side of the webpage [40]. Moreover, during browsing and searching text-based websites, they follow Jakob Nielsen's "F-pattern" [40]. One of the web behaviors is when users focus more on the left side of the webpage than on the right side through reading [40]. Considering the "F-pattern," Nielsen [41] stated that because it takes an effort to read an on-screen text, therefore users just scan and read the parts which are interesting for them.

TABLE II. HYPOTHESES ABOUT ADVERTISING FEATURE FOR THE LOW/HIGH CD TASK AND IMPACT OF GENDER

No.	*Hypothesis*	
H1A/H1B	When people are engaged with a low/high CD task,	Ad with fewer words will attract a) earlier, b) greater, and c) more frequent visual attention than the ad with more words.
H2A/H2B		Ad with underlined words will attract a) earlier, b) greater, and c) more frequent visual attention than the ad with no underlined words.
H3A/H3B		Ad with words colored in warm colors (red is discussed in this research) will attract a) earlier, b) greater, and c) more frequent visual attention compared to the ad with no words in warm colors.
H4A/H4B		Personalized ad (including participants name) will attract a) earlier, b) greater, and c) more frequent visual attention compared to the non-personalized ad.
H5A/H5B		Personalized ads with underlined words will attract a) earlier, b) greater, and c) more frequent visual attention than the non-personalized and non-underlined words ad.
H6A/H6B		The ad containing the photograph of a person while drinking tea will attract a) earlier, b) greater, and c) more frequent visual attention than the ad without this photograph.
H7A/H7B		The ad with a darker background will attract a) earlier, b) greater, and c) more frequent visual attention than the ad with a brighter background.
H8A/H8B		The ad that is located on the right side of the webpage will attract a) earlier, b) greater, and c) more frequent visual attention than the ad located on the left side of the webpage.
H9A/H9B		The ad with a larger size will attract a) earlier, b) greater, and c) more frequent visual attention than the smaller ad.
H10A	Gender affects users' attention in the aspect of a) Time To the First Fixation (*TTFF*), b) Total Fixation Duration (*TFD*), and c) Fixation Count (*FC*) for the low CD task.	
H10B	Gender affects users' attention in the aspect of a) *TTFF*, b) *TFD*, and c) *FC* for the high CD task.	
H11	CD-level of the task impacts on users' attention to the ads varying in the nine mentioned features, in terms of a) *TTFF*, b) *TFD*, and c) *FC*.	

In the last experiment of the current research, the best English free website of classified advertising website in the USA for 2018, according to ads2020 website is displayed to the users with a right to left direction mother tongue language (Farsi) and presents different results by analyzing their gaze plots. Since these participants have an unconscious right to the left direction for reading, this research claims a different starting viewpoint. The eye-tracking technique reveals this difference with gaze plots. The hypothesis of this claim is as follows:

Hypothesis H12: Users with right to left direction mother tongue language, first look at the right side of the English advertising webpage.

Since users pay attention to the website design elements (including text) that are most important to them, authors of websites and designers could apply models of visual attention to improve the design of their website [5]. The eye-tracking technique provides users' interaction with interfaces, including the most and first viewed elements data. Besides, this data helps them study the effect of changes in design to achieve their favorable users' gaze patterns [42]. Concerning these results, a better user experience is provided for the users.

2-4. USABILITY IMPORTANCE

Human-computer interaction is the discipline concerned with designing, evaluating, and implementing interactive computing systems for human use and studying significant phenomena surrounding them [15]. The core activity in human-computer interaction studies over the past fifteen years is to develop effective usability inspection techniques [16].

International Organization for Standardization (ISO) describes usability as the extent to which specified users can use a product to achieve the specific goals with effectiveness, efficiency, and satisfaction in a determined context of use [17]. For the new technologies, usability metrics are noteworthy in providing positive user satisfaction [6]. Therefore, since many websites are available to users, they disapprove of bad usability. Eye-tracking is a technique for evaluating usability, which records eye movements as the user is looking at the stimulus. This study uses this tool to examine different advertising features to attain the quickest, longest, and most frequent attention among users' diverse characteristics.

2-5. EYE-TRACKING TECHNIQUE IN HUMAN-COMPUTER INTERACTION

Eyes have many micro-movements per second that can reach a few pixels. A nearly motionless moment of an eye is called a fixation. An eye-tracking software recognizes a fixation from data gathered by the eye-tracker. Examples of visualizations metrics are heat maps, which display the intensity of looking at the stimuli, and gaze plots that show individuals' fixation succession. Data collected by the eye-tracking technique form users' visual behavior usability analysis [43]. This data is reliable even when usability-testing methods are not satisfying [44]. For instance, eye movements are able tformeal distraction even though tasks are accomplished. Moreover, since eye tracking provides insights into problem-solving, reasoning, mental imagery, and search strategies, cognitive psychology offers a rich background in usability. Early studies using eye-tracking began before computers, as we now know them, were introduced [43].

This research uses eye-tracking software to measure participants' visual attention to the ads. The samples of users' pupils' recordings by the eye-tracker provide the visual time spent on the ad as the Area of Interest (AOI) [2]. There is a relationship between eye movement metrics and cognitive processes or usability problems. For example, TTFF on target is good (if short) or bad (if long) attention-getting property [43,45], higher fixations per AOI points out a more noticeable or more important element/area [43,46,47], and fixation duration exposes difficulty in extracting information or more engaging, voluntary (>320 ms) and involuntary (<240 ms) fixations [47,48]. Test facilitators avoid enormous error, such as interrupting, while the user is silent for a long time through using eye-tracking technology. Since inappropriately or too much intrusion of the participant is a critical problem of usability professionals, this technology is highly advantageous [40]. Also, other researchers have established this method [49-56].

3. METHOD

The greatest achievement of this study obtained by the eye-tracking technique is first, presenting features of ads that promote its function after intelligently indicating the CD level of users' tasks. Secondly, analyzing the visual behavior of people with the right to left direction mother tongue language while looking at an English advertising website.

To examine the features of an ad, this study provides tables that consist of values of measures derived from the eye-tracker device. The experiment was implemented in a cognitive laboratory equipped with a desktop computer, Tobii TX300 eye-tracker, and an observatory room for the researcher. The participants were told: "In this study, we will be observing what you look at as you use the website. Thus we will use technology to track your eye movements."

The SVM algorithm is applied to users' gaze pattern dataset by Weka 3.8 machine learning software to classify CD levels of tasks.

3-1. PARTICIPANTS

In the first phase, ten graduate and undergraduate students from the University of Tehran, consisting of 5 males and five females, were studied. In the second phase of the experiment, 30 graduate and undergraduate students were the research participants from the University of Tehran. They were between 18 and 37 (Mean =23.46) years old. Gender distribution was 53% female and 47%, male. All users' visual abilities were treated to be healthy or normal.

3-2. PROCEDURE

All participants registered for a 30 minutes experiment. The researcher collected each persons' information after they entered the lab and then explained the investigation process. Furthermore, the researcher assigned a sheet explaining the whole process, including the calibration method and commitment to not interfere with the results. First, participants were told to look at each of the pages with the goal of free-viewing (low CD task) for 25 seconds. Secondly, the researcher declared them to answer the researchers' questions from the text (high CD task). After the researcher informed the participants, she started the calibration process and then uploaded the websites consisting of newsletters and banner ads. Then, participants read the articles at their own pace as they were completing the given task. Finally, for the third task, they told the participant to look at an advertising website with the goal of free viewing.

3-3. STIMULUS MATERIAL

According to a published news article, the researcher composed an online news story named "knowing the most beautiful waterfalls in our country" in a webpage format. Each article explains various features of a waterfall, including height, history, location, and photographs. An interesting topic was chosen to involve adult participants. Since the participants were all Asian and tea is their most common drink, this study selects it for the advertising product. Moreover, in each article, ten versions of the banner ad exist. To prevent any previous experiences or exposure to the brand from interfering with the experiment, this study exploits an imaginary brand ("Tara tea"). Furthermore, the location of each ad is next to the created webpage with a "top news" title. In the last experiment, the researcher displayed Craiglist website, the best English free classified advertising website in the USA for 2018, according to the ads2020 website to the users.

3-4. MANIPULATION

This paper studies ten ads with various attributes. Features of the first phase include the number of words, underlining or coloring them with warm colors (this research studies red), and personalizing (inserting each participant's full name). Ads announce a 10% discount for the students who join the opening in the coffee shop (e.g., Ms. Tina Raad, you are invited to the opening of Tara tea. Do not miss our 10% discount for the students.). In the primitive non-personalized ad state, participants view a more general sentence ("A warm morning with Tara tea"). Features of the second phase also consist of both underlining and personalizing, placing a picture of someone while drinking tea, darkening the background color, changing the location (this research studies placements in left and right side of the webpage) and changing the size (at first this research studies 168px (weight) × 250px (height) and then 2) 228px (weight)×340px (height) as the larger attribute.). In manipulating the complexity of the task, the high CD ones are fact-findings, and the low CD is free-viewings. In the high CD task, the researcher requires participants to find the answers to the questions that she asks them from the news article [2, 57].

3-5. MEASURES

1) ELAPSED TIME TO THE FIRST FIXATION ON THE AOI (TTFF): These metric measures the time passed since the system loads the news article until participants' first fixation on the area of interest, which is the ad by the eye-tracking software [3]. The eye-tracker captures samples at the rate of 60 Hz, or 60 samples per second, to collect data.

2) TOTAL VISUAL ATTENTION PAID TO AOI (TFD): This metric measures the sum of the duration for all fixations within an AOI (e.g., ad) in seconds [3]. The eye-tracker software specifies it with a total fixation duration metric.

3) RECURRENCE OF ATTENTION TO THE AOI (FC): The eye-tracker software defines this metric as fixation count, which counts the number of times a participant fixates on the AOI (e.g., ad) [3]. If during the recording, a participant leaves, and returns to the same media element, then the new fixations on the media will be included in the calculations of the metric. If the participant has not fixated on the ad at the end of the recording, the eye-tracking software does not compute the fixation count value. It does not include the recording in the descriptive statistics calculations.

4) GAZE PLOT: This visual metric contains information on an individual level about the order and position of gaze points. The eye-tracker specifies this metric for the whole duration of the experiment.

4. RESULTS

In this section, the results of the first and second phases and tasks classification are provided.

4-1. FIRST PHASE

The first phase examines participants in a free–viewing task for the following ad features of Tara tea: F0: basic, F1: fewer words, F2: underlined words, F3: warm-colored words, and F4: personalized. These features are compared based on their mean amount in TTFF, TFD, and FC metrics in Fig. 1. In the chart of Fig. 1, personalizing has the highest amount in TFD and FC. However, the ad with underlined words has the least amount in TTFF. Therefore, on average, the personalization of the ad promotes the duration and frequency of users' attraction, whereas underlining the word makes it catchier at first sight for them. Additional study on the situational effectiveness of this experiment, for example, in other types of websites and ads, is needed. The second phase studies the combination of these measures and four more features of the ad.

4-2. TESTING HYPOTHESES OF TABLE I

In the second phase, users are studied at first while accomplishing the low CD task and then the high CD task for the following features of the ad, which are respectively obtained from the hypotheses in Table II: F0: basic, F1: fewer words, F2: underlined words, F3: warm-colored words, F4: personalized, F5: both underlined words and personalized, F6: containing a photograph of a person while drinking tea, F7: darker background, F8: located in the right side of the webpage and F9: larger size. To evaluate H1A, … H9A, H1B, … and H9B hypotheses, each of the TTFF, TFD, and FC metrics are assessed in both low and high CD tasks with statistical approaches a one-way ANOVA test, Post Hoc test, and Tukey method. One-way ANOVA test specifies whether there are any significant differences among features. If such differences exist, Post Hoc and Tukey tests decipher which features have caused them. They compare these features by providing p-values and their Mean Difference (MD) for every two features, such as Fi and Fj (MD(i,j)). If $MD(i,j) \geq 0$, then Fi has a higher amount than Fj in the related metric.

In this research, p-value (p) determines whether any of the differences between the means are statistically significant, and the significance level denoted as α is 0.05. In assessing the null hypothesis of a statistical test, p is compared with α. If $p>\alpha$, the null hypothesis is approved. The differences between the means are not statistically significant, whereas if $p<\alpha$, then the differences between the means are considered statistically significant, and the null hypothesis is rejected. Moreover, F is a symbol for combining variance between sample means, the variance within the samples, and the sample sizes. If F has a large value, equal means are less likely. Additionally, the Degree of Freedom (df) equals one less than sample sizes. Thus, the value of it for ten ads in one-way ANOVA is 9. Also, Standard Error (Std. Error) measures the accuracy as a sample represents a population. In the second phase of this research, three one-way ANOVA tests are applied to the participants' TTFF, TFD, FC data. Their results for low and high CD tasks are shown in Table III. This table does not contain data of the features that do not have at least one significant difference with F0 for the three metrics. The effects of the features on each metric for both tasks are discussed in the following. Additional study on the situational effectiveness of this experiment, for example, in other types of websites and ads, is needed.

TABLE III. LOW AND HIGH CD TASKS

	Low CD task			High CD task		
	Df	*F*	*P*	*df*	*F*	*P*
TTFF	9	1.54	0.13	9	4.47	0.00
TFD	9	7.37	0.00	9	3.01	0.00
FC	9	4.45	0.00	9	2.78	0.00

1) *TESTING H1A, …, AND H9A*

Firstly, the function of ads in attracting users' sooner is explained. According to Table III, there are no significant differences among the nine features in the aspect of TTFF (p=0.13, p>0.05, F=1.54, df=9). In other words, no ad is more attractive than the other at first sight in the low CD task. Consequently, the first parts of H1A, …, H9A are rejected. Secondly, the function of ads in attracting users for a longer duration is argued. Based on Table III, at least two features have a significant difference in TFD (p=0.00, p<0.05, F=7.37, df=9). Based on Table IV, p is less than 0.05 for all features except F4 in comparison with F5 (p=0.71, p>0.05), and all of them have positive MDs. In other words, underlining words and personalizing the ad simultaneously leads to longer attraction. Therefore, the second part of H5A is accepted, and the second part of other hypotheses except for H4A are rejected. Moreover, p is less than 0.05 for F4 in comparing with F0 and has a positive MD (p=0.00, p<0.05, MD(4,0)=3.12). Therefore, personalizing the ad improves the duration that users are attracted to it. In consequence, the second part of H4A is accepted. It should be noted that although F5 does not have a significant difference in TFD with F4, it has higher MD than F4 in comparison with F0 (MD(4,0)=3.12, MD(5,0)=4.78, MD(5,0)>MD(4,0)). Consequently, the personalized ad with underlined words has the highest length of visual attraction in the low CD task. Finally, the function of ads in attracting users more frequent is described. According to Table III, there is a significant difference between at least two features in the aspect of FC (p=0.00, p<0.05, F=4.45). To determine them, Post Hoc and Tukey tests are applied. Their results are shown in Table IV. Based on Table IV, F5 has the highest amount of FC (p=0.00, p<0.05, MD(5,0)=8.02). In other words, personalizing and underlining words at the same time leads to users' higher frequency of attention. On the other hand, F4 and F5 do not have a significant difference with each other in terms of FC, but F5 has a higher mean difference (MD(5,4)=2.13), which means that underlining the personalized ad increases the number of times a user gets attracted to it.

As a result, the third part of H5A is approved, and the third part of the other hypotheses are rejected. In brief, a personalized ad with underlined words catches users more often than the non-personalized or only personalized ones. The heat map of F5 for one of the participants is shown in Fig. 2. According to Fig. 2, a considerable attraction is focused on the ad demonstrated with red color. Furthermore, the values of Std. Error in Table IV for the argued features are less than three, which almost reveals accuracy in the related data to it. Additionally, Fig. 4 briefly demonstrates the explained evaluations of hypotheses in Table II. In Fig. 4, red rectangles show worsens states, and green rectangles show the improved ones compared to the primary webpage. Additional study on the situational effectiveness of this experiment, for example, in other types of websites and ads, is needed.

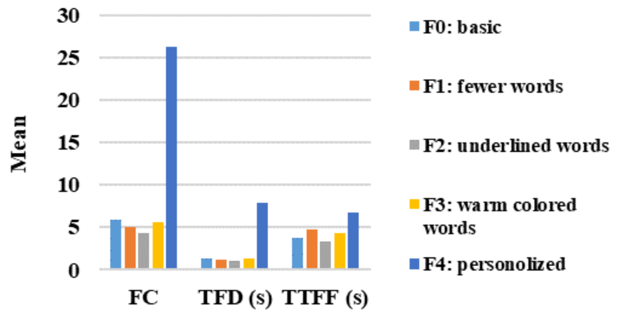

Fig. 1. Comparisons of the features in the first phase by the eye-tracking software

Fig. 2. Heat map of *F5* in the low CD task for one of the users by the eye-tracking software

2) *TESTING HIB (I=1, …, 9)*

Same as Task 1, three one-way ANOVA tests evaluate H1B, …, and H9B hypotheses in Table II for the high CD task. The results of these assessments are shown in Table III. Based on Table III, there are at least two features that have significant differences with each other in TTFF (p=0.00, p<0.05, F=4.47, df=9), TFD (p=0.00, p<0.05, F=3.01, df=9) and FC (p=0.00, p<0.05, F=2.78, df=9). To decipher these features, Post Hoc and Tukey tests are applied to data of high CD tasks, and their results are demonstrated in Table IV. According to Table IV, all three metrics are discussed in the following:

Firstly, in the aspect of TTFF, F7 has a significant difference in comparing with F0 (p=0.03, p<0.05, MD(7,0)=8.16) and F1 (p=0.03, $p<0.05$, MD(7,1)=9.05). Also, F9 has a significant difference in comparing with F0 (p=0.00, $p<0.05$, MD(9,0)=10.21) and F1 (p=0.00, $p<0.05$, MD(9,1)=11.10). Since all MD(7,0), MD(7,1), MD(9,0), and MD(9,1) have positive values, users are later attracted to the ad with a darker background or larger size. Therefore, the first parts of H7B and H9B are rejected. Secondly, from TFD perspective, F6 has a significant difference with F7 (p=0.03, $p<0.05$, MD(6,7)=-1.28, MD(6,7)<0). Consequently, adding the photograph of a person while drinking tea to the ad induces shorter attention than setting a darker background for it.

Since no feature has a significant difference with F0, the second parts of H1B, H2B, …, H9B are rejected. At last, in terms of FC, F7 has a significant difference with F0 (p=0.04, $p<0.05$, MD(7,0)=3.79), F1 (p=0.01, $p<0.05$, MD(7,1)=4.66) and F6 (p=0.01, $p<0.05$, MD(7,6)=4.73). In other words, the ad with a darker background attracts users more often than the one that has fewer words or the photograph of someone while drinking tea. In brief, however, darkening the ad grabs participants' attention later, they pay more frequent and longer attention to it than the other ads. The heat map of F7 for all participants is shown in Fig. 3. Based on Fig. 3, users focused on the ad and mostly the answer to the question they have been asked for. Moreover, in contrast with many advertisers' opinions, the larger ad does not improve its function for any of the three so-called metrics—furthermore, the values of Std. Error in Table IV for the discussed features is less than three, which shows good accuracy in the related data. Also, Fig. 4 shortly describes the interpretations of the hypotheses in Table II. Additional study on the situational effectiveness of this experiment, for example, in other types of websites and ads, is needed.

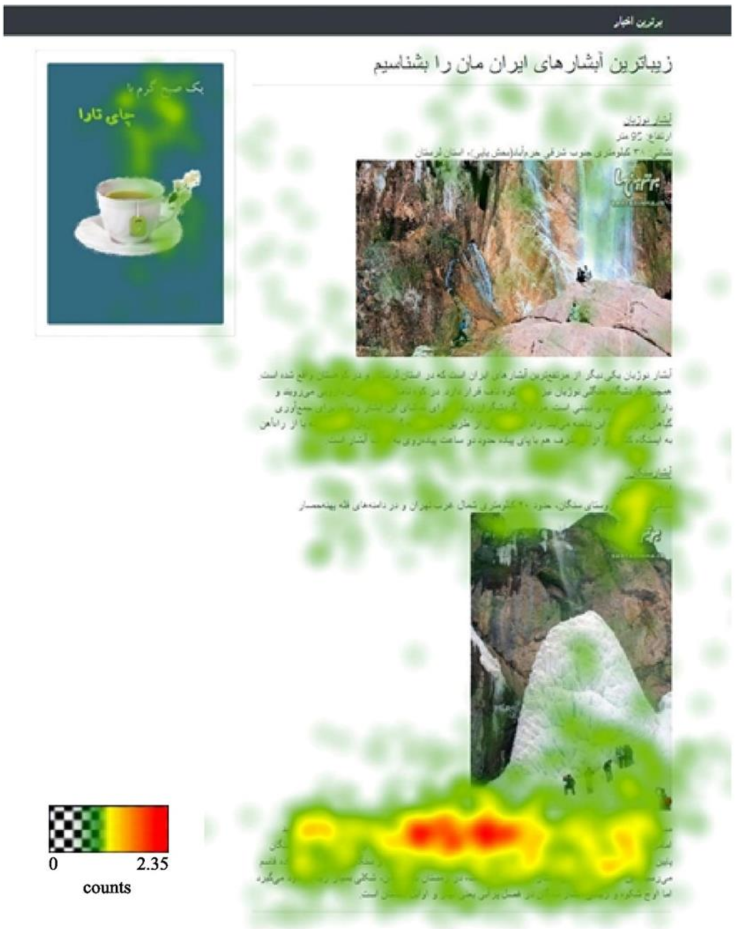

Fig. 3. Heat map of *F7* in high CD task for all users by the eye-tracking software

3) TESTING HYPOTHESIS 10A

This study uses the independent t-test to reveal the effect of gender on participants' attention in the low CD task for TTFF, TFD, and FC. An independent t-test has F, t, df, p, MD, and Std. Error Difference variables are explained in the following. The ratio of the difference among the means of the data related to each gender and its difference is expressed with t. A higher value of t indicates a large difference among males and females, and a smaller amount shows more similarity. Moreover, a Mean Difference (MD) shows the difference between data for males and females in the three metrics. Also, Standard Error Difference (Std. Error Difference) between the means serves the same purpose as a standard deviation within the context of an independent t-test. A standard deviation applies to one group of data, but a standard error of the difference among the means applies to two groups of data. Lower variability around the group means leads to lower Std. Error Difference. Three independent t-tests are applied on data of TTFF, TFD, and FC metrics in the low CD task. The results of these tests are shown in Table V and are discussed in the following.

First, there is no significant difference between males and females in terms of TTFF (all p-values are greater than 0.05). Thus, on average, males and females function in the same way in terms of the elapsed time to fixate on the ten ads for the first time. In consequence, the first part of H10A is rejected. Secondly, in the aspect of TFD, F3 and F9 have significant differences between males and females (F3: p=0.01, $p<0.05$, F=2.77, t=2.93, MD=2.62, Std. Error Difference=0.89 and F9: p=0.01, $p<0.05$, F=8.59, t=2.83, MD=0.79, Std. Error Difference=0.27). To mathematically describe the effect of gender by formulas, linear regression models are used. Their results are provided in Table VI. This table includes models consisting of constants and G variables, unstandardized coefficients with their Std. Error, t, and p. According to Table VI, no coefficient in the formula of TFD is zero (all p-values are less than 0.05). Besides, column B provides values of coefficients in each row for the so-called formulas. Thus, the constant and coefficient values of variable G are respectively 3.42 and -2.62 for F3 and 1.24 and -0.79 for F9. As a result, their formulas are as (1) and (2):

G={0: female, 1: male}

TFD:

$$F3=3.42-2.62\times G \qquad (1)$$

$$F9=1.24-0.79\times G \qquad (2)$$

Based on (1) and (2), on average, males spend 2.62 seconds on the ad with words colored in red and 0.79 seconds on the larger one, less than females in this task. Therefore, larger and warm-colored sentences attract males for a shorter time duration in the low CD task. Finally, in sense of FC these two features also have significant differences among males and females (F3: p=0.03, $p<0.05$, F=1.49, t=2.39, MD=7.71, Std. Error Difference=3.22 and F9: p=0.01, $p<0.05$, F=14.43, t=2.73, MD=3.42, Std. Error Difference=1.25). Thus, the linear

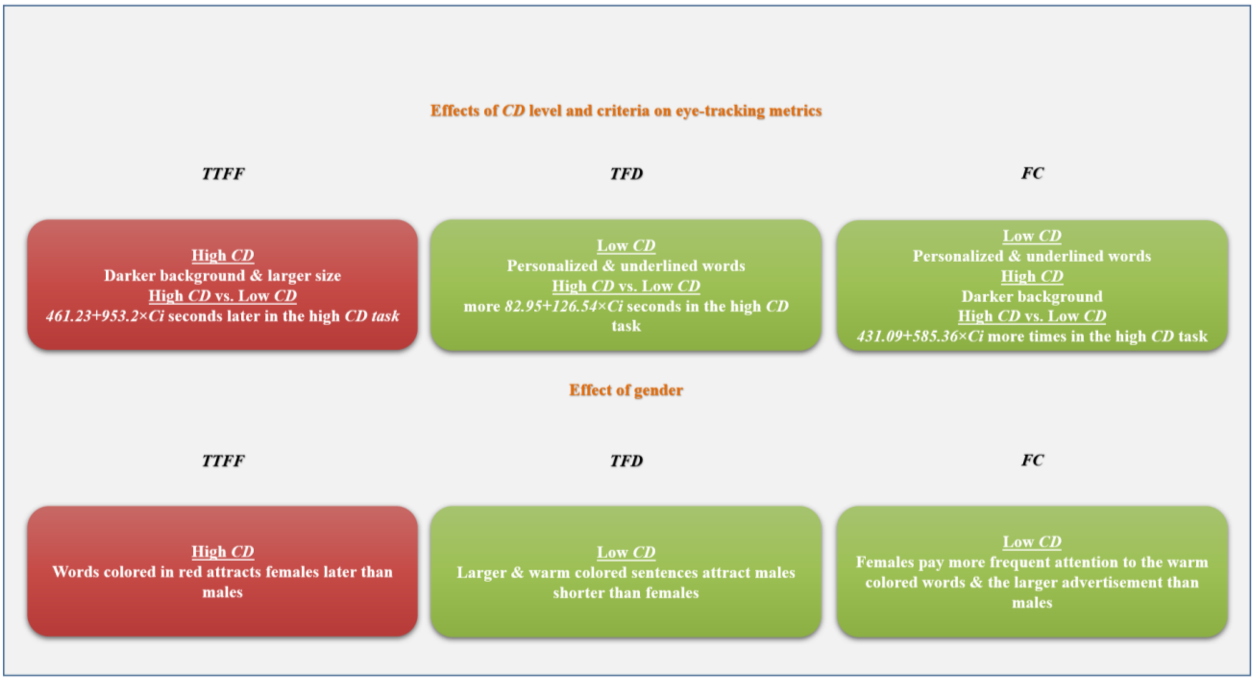

Fig. 4. Effects of CD level, features, and gender on users' attention based on eye-tracking metrics

TABLE IV. FEATURES MEASUREMENT FOR *TTFF*, *TFD* AND *FC* WITH TUKEY HSD TEST IN THE LOW AND HIGH CD TASKS

Low CD task						**High CD task**					
Measures	***Fi***	***Fj***	***MD(i,j)***	***Std. Error***	***p***	***Measures***	***Fi***	***Fj***	***MD(i,j)***	*Std. Error*	*p*
TFD	*F0*	*F1*	-0.48	0.91	1.00	*TTFF*	*F7*	*F0*	8.16	2.43	0.03
		F2	0.57	0.87	1.00			*F1*	9.05	2.71	0.03
		F3	-0.75	0.89	0.99			*F2*	-2.45	4.26	1.00
		F4	-3.12	0.77	0.00			*F3*	2.17	2.96	0.99
		F5	-4.78	0.89	0.00			*F4*	8.87	3.64	0.31
		F6	-0.17	0.79	1.00			*F5*	-0.36	3.05	1.00
		F7	-0.12	0.78	1.00			*F6*	1.33	2.71	1.00
		F8	0.37	0.75	1.00			*F8*	6.02	2.71	0.45
		F9	0.43	0.81	1.00			*F9*	-2.05	2.63	0.99
FC	*F0*	*F1*	-0.24	2.15	1.00	*TFD*	*F6*	*F0*	-0.17	0.35	1.00
		F2	1.66	2.06	0.99			*F1*	-0.06	0.39	1.00
		F3	-1.95	2.10	0.99			*F2*	-0.31	0.61	1.00
		F4	-5.89	1.84	0.05			*F3*	-0.26	0.43	1.00
		F5	-8.02	2.10	0.00			*F4*	-1.00	0.52	0.66
		F6	-1.04	1.87	1.00			*F5*	-1.38	0.44	0.06
		F7	-0.57	1.84	1.00			*F7*	-1.28	0.38	0.03
		F8	1.10	1.77	1.00			*F8*	-0.31	0.39	0.99
		F9	1.22	1.92	1.00			*F9*	-0.85	0.38	0.45
TFD	*F5*	*F0*	4.78	0.89	0.00	*TTFF*	*F9*	*F0*	10.21	2.43	0.00
		F1	4.29	1.02	0.00			*F1*	11.10	2.71	0.00
		F2	5.35	0.99	0.00			*F2*	-0.40	4.26	1.00
		F3	4.02	1.00	0.00			*F3*	4.22	2.96	0.91
		F4	1.65	0.90	0.71			*F4*	10.92	3.64	0.09
		F6	4.61	0.92	0.00			*F5*	1.68	3.05	1.00
		F7	4.65	0.91	0.00			*F6*	3.39	2.71	0.96
		F8	5.15	0.88	0.00			*F7*	2.05	2.63	0.99
		F9	5.21	0.94	0.00			*F8*	8.07	2.71	0.09
FC	*F5*	*F0*	8.02	2.10	0.00	*FC*	*F7*	*F0*	3.79	1.16	0.04
		F1	7.78	2.41	0.04			*F1*	4.66	1.30	0.01
		F2	9.69	2.32	0.00			*F2*	3.50	2.04	0.78
		F3	6.07	2.36	0.24			*F3*	3.72	1.42	0.21
		F4	2.13	2.14	0.99			*F4*	2.33	1.74	0.94
		F6	6.97	2.16	0.04			*F5*	1.20	1.46	0.99
		F7	7.44	2.14	0.02			*F6*	4.73	1.30	0.01
		F8	9.12	2.07	0.00			*F8*	3.20	1.30	0.30
		F9	9.25	2.20	0.00			*F0*	3.79	1.16	0.04

regression model attains its formulas as (3) and (4):

FC:

$$F3=11.14-7.71\times G \qquad (3)$$

$$F9=5.54-3.42\times G \qquad (4)$$

According to (3) and (4), males fixate 7.71 times on the ad with warm-colored words and 3.42 times on the larger one less than females. Consequently, larger ads and warm-colored sentences attract males shorter and less frequent than females in low CD tasks. Therefore, the visual attention of males and females differs only in the color of words and the size of the ads in the aspect of the three so-called metrics in this task. As a result, the second and the third parts of H10A for F3 and F9 are approved, and the others are rejected. Fig. 4 briefly explains these notes.

4) *TESTING HYPOTHESIS 10B*

Three independent t-tests are applied to the data related to TTFF, TFD, and FC in the high CD task. Table V shows their results. According to Table V, there is no significant difference between males and females in terms of TFD and FC (all p-values are greater than 0.05), so the second and third parts of H10B are rejected. Based on Table V, F3 has a significant difference between males and females in the aspect of TTFF ($p=0.01$, $p<0.05$, $F=15.68$, $t=3.08$, $MD=10.78$, Std. Error Difference=3.49). To mathematically demonstrate the effect of gender on TTFF, the linear regression model is used to provide a formula. Table VII shows the results of this model. Same as the explained approach in Section 3, the attained formula for F3 is as (5):

TTFF:

$$F3=11.63-10.78\times G \qquad (5)$$

Thus, on average, the ad with words colored in red attracts males 10.78 seconds sooner than females in the high CD task. Briefly, males and females differ only in the color of words feature for the TTFF metric in the high CD task. Consequently, the first part of H10B is accepted for F3. Fig. 4 summarizes these explanations. Additional study on the situational effectiveness of this experiment, for example, in other types of websites and ads, is needed.

5) *TESTING HYPOTHESIS 11*

This paper studies the common features of an ad in two tasks with low and high CD Levels (CDL). To test whether CD

TABLE V. IMPACT OF GENDER ON THE FEATURES IN THE LOW AND HIGH CD TASKS

		low CD task						***high CD task***					
Measures	***Fj***	***F***	***t***	***df***	***p***	***MD***	***Std. Error Difference***	***F***	***t***	***df***	***p***	***MD***	***Std. Error Difference***
TTFF	*F0*	17.14	-2.03	9.79	0.07	-3.74	1.84	4.09	-1.04	22	0.30	-2.31	2.21
	F1	2.11	0.75	11	0.46	2.90	3.82	4.27	-0.91	13	0.37	-2.60	2.86
	F2	0.52	0.50	13	0.62	0.80	1.58	.	-0.99	2	0.42	-10.65	10.73
	F3	0.10	-1.09	12	0.29	-4.92	4.49	15.68	3.08	7.20	0.01	10.78	3.49
	F4	0.44	-0.94	21	0.35	-2.20	2.32	14.77	-1.00	2.00	0.42	-3.14	3.13
	F5	0.45	-0.09	12	0.92	-0.38	4.29	0.00	1.53	8	0.16	8.10	5.28
	F6	0.01	0.40	19	0.69	1.19	2.98	0.13	0.31	13	0.75	1.49	4.71
	F7	0.03	0.17	20	0.86	0.55	3.14	1.12	-0.48	15	0.63	-2.17	4.51
	F8	0.30	-0.18	24	0.85	-0.38	2.10	0.02	-0.25	13	0.80	-1.14	4.42
	F9	0.00	-0.81	17	0.42	-2.83	3.47	0.40	-0.15	15	0.88	-0.59	3.89
TFD	*F0*	4.11	0.90	22	0.37	0.49	0.54	0.01	-0.71	22	0.48	-0.10	0.15
	F1	0.00	-0.01	11	0.98	-0.01	1.37	4.61	1.36	13	0.19	0.20	0.14
	F2	1.20	0.26	13	0.79	0.09	0.34	.	0.87	2	0.47	0.86	0.98
	F3	2.77	2.93	12	0.01	2.62	0.89	0.30	0.17	9	0.86	0.10	0.56
	F4	2.65	0.41	21	0.68	0.74	1.80	3.25	-1.34	4	0.24	-1.61	1.19
	F5	0.79	-0.52	12	0.61	-2.01	3.86	3.67	-0.65	8	0.52	-0.94	1.43
	F6	2.30	0.23	19	0.81	0.08	0.38	0.01	0.16	13	0.87	0.03	0.20
	F7	0.52	0.59	20	0.55	0.42	0.71	0.23	0.51	15	0.61	0.35	0.69
	F8	0.20	-1.44	24	0.16	-0.42	0.29	4.49	-1.04	13	0.31	-0.70	0.67
	F9	8.59	2.83	12.71	0.01	0.79	0.27	0.11	0.29	15	0.77	0.17	0.61
FC	*F0*	4.25	1.23	22	0.22	2.45	1.98	0.02	-0.16	22	0.87	-0.11	0.73
	F1	0.28	0.25	11	0.80	0.76	3.05	0.23	-0.17	13	0.86	-0.08	0.52
	F2	0.37	0.10	13	0.91	0.17	1.68	.	0.89	2	0.46	3.33	3.71
	F3	1.49	2.39	12	0.03	7.71	3.22	0.49	-0.03	9	0.97	-0.08	2.34
	F4	1.89	-0.14	20	0.88	-0.68	4.83	1.61	-1.27	4	0.27	-4.00	3.14
	F5	0.10	-0.45	12	0.65	-2.95	6.50	0.02	-0.19	8	0.85	-0.80	4.19
	F6	4.06	1.01	19	0.32	1.56	1.54	0.51	0.48	13	0.63	0.50	1.03
	F7	2.04	1.33	20	0.19	2.57	1.93	1.80	0.96	15	0.34	2.31	2.39
	F8	0.62	-0.83	24	0.41	-0.92	1.10	5.25	-0.69	7.753	0.50	-1.50	2.15
	F9	14.43	2.73	11.15	0.01	3.42	1.25	0.02	0.23	15	0.81	0.54	2.31

of a task affects users' attention to the ads in TTFF, TFD, and FC metrics, two-way ANOVA tests are used, and their results are shown in Table VIII. In this table, each metric has a Source for itself that includes the independent variables, their interaction, and an intercept. Additionally, the Type III Sum of Squares (SS), which is used for values of coefficients in equations, when an interaction is present between the independent variables, is defined for each metric. According to Table VIII, all p-values are less than 0.05, so there are significant differences among features in high and low CD levels. Thus, all coefficients of the equations are non-zero and significant. The achieved equations are as (6), (7) and (8) for each metric:

$$TTFF=10989.86+2133.25\times Fi+461.23\times CDL+953.2\times Fi\times CDL \quad (6)$$

$$TFD=665.35+237.31\times Fi+82.95\times CDL+126.54\times Fi\times CDL \quad (7)$$

$$FC=7722.7+851.69\times Fi+431.09\times CDL+585.36\times Fi\times CDL \quad (8)$$

where CDL is respectively defined as 0 and 1 for the low and high CD tasks. Also, Fi is defined as i for the nine features (i=0,…,9), and F0 is the basic ad.

Based on the above formulas, the CD level of the task affects users' visual behavior for the studied features in all three measures. Based on (6), in combining impacts of features and CD level of the task on users' visual attention, high CD task attracts users 461.23+953.2×Fi seconds later than the low CD tasks. According to (7) and (8), in comparing with the free-viewing state, users spend more 82.95+126.54×Fi seconds and 431.09+585.36×Fi times on the ad with the Fi characteristic in the high CD task. Fig. 4 briefly provides these results.

6) TASKS CLASSIFICATION

In the previous sections, the relation among users' visual behavior, types of tasks, and ad is recognized. Consequently, placing the best ad on the website after intelligently realizing users' type of task is essential. The eye-tracking technique provides gaze plots of users, which determines the location of each gaze point from the start to the end of the experiment. Each web page is divided into nine equal areas: A, B, C, D, E, F, G, H, and I. Fig. 5 shows this division. To classify the tasks, each sequence of the first five gaze points for every user is labeled as the area placed on (e.g., A, E, C, F, F). A dataset of all users' first five gaze points is made to classify the tasks, consisting of 60 sequences.

Regarding the size and sequential format of this dataset, the SVM algorithm is used for classification. Table IX shows the results for different properties of this algorithm. Moreover, it reveals a decision stump tree and J. 48 algorithms that do not provide results as good as SVM. According to Table IX, the highest accuracy is 80% percent. The properties of these algorithms are set as the polynomial kernel, right probability estimates, and nu-SVC for the SVM type. Consequently, this paper intelligently presents ads to users by figuring out their first five gaze points with high accuracy. Additional study on the situational effectiveness of this experiment, for example, in other types of websites and ads, is required.

4-3. TESTING HYPOTHESIS 12

The eye-tracking software provides gaze plots of users, which determines their starting points of gazing. Fig. 6 shows all participants' results for up, down, middle, left, and right sides of the webpage. Based on this Figure, users with the right to left direction mother tongue languages' gazing start point for English language websites are descending order middle, right, down, left, and up. Additionally, based on the previous studies, users appeared to follow Jakob Nielsen's F-pattern [25] when browsing and searching text-based websites. In contrast with Jakob Nielsen's F-pattern [40], these users have a T shape pattern. As a result, contradicting Jacob F-pattern, people with the right to left direction mother tongue language at first look at the middle of the English website on average. The gaze plot of all participants for the English language advertising website is shown in Fig. 7. Each participant's sequence of gaze points is shown with circles with a specific color and numbers in this figure. These numbers determine the order of gaze points, and the size of each circle indicates the length of the related fixation. According to Fig. 6, most users' first gaze point is in the middle of the page. This conclusion helps website designers be aware of the visual gaze patterns of the so-called users and improve their designs. It is also considerable that many English websites are viewed by people with a right to left mother tongue language, and F-pattern-based designs are not the most efficient designs for them. Additional study on situational effectiveness of this experiment, for example, is needed in other types of websites.

5. DISCUSSION AND CONCLUSION

This paper studies nine features of an online ad for Tara tea in two phases. The first phase analyzes four features, including fewer, underlined, and warm-colored words and personalizing the ad. Evaluations of these features show that inserting individuals' names promotes users' total duration and frequency of attention to the ad. Although, underlining the words has the best function in catching users' attention at first glance. To

Fig. 5. Labels of area division for users' gaze pattern sequence

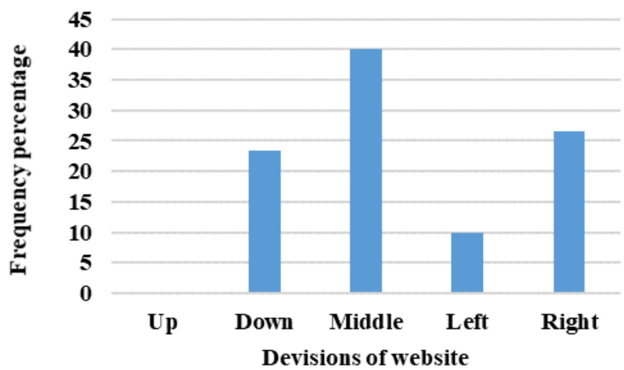

Fig. 6. The starting point of gazing for advertising website

complete this experiment, the second phase studies the simultaneous combination of underlining words and personalizing features and eight other characteristics of the ad as the following: fewer, underlined, and warm-colored words, personalizing, containing a photograph of someone while drinking tea, darkening the background, placing the ad in the right side of the webpage and larger size.

Previous literature suggests the CD level of a task impacts users' visual attention to the ads [2]. This paper studies advertising features in both high and low CD tasks, and the findings are as follows:

- In the low CD state, personalizing the ad and underlining its words at the same time has the best function in attracting users for the longest duration and more frequent attention among all features. This new finding emphasizes the importance of combining personalization and underlining the words features of ads to improve users' attention.

- In the high CD tasks, ads with a larger size or a darker background were less attractive to catch users' attention at first glance. Also, compatible with previous research contending that pictures are more attractive to users than texts themselves, this study shows that inserting a picture of someone while drinking tea attracts users for a longer time duration than darkening the background. Although darkening the background is not successful in promoting grabbing users' attention sooner and longer, it attracts them more frequently. This new finding implies that unlike inserting a picture of someone while drinking tea, the background color is not very successful in lengthening users' attention.

- Opposed to the common belief that a larger ad increases users' visual attention, it does not cause a specific improvement in any task.

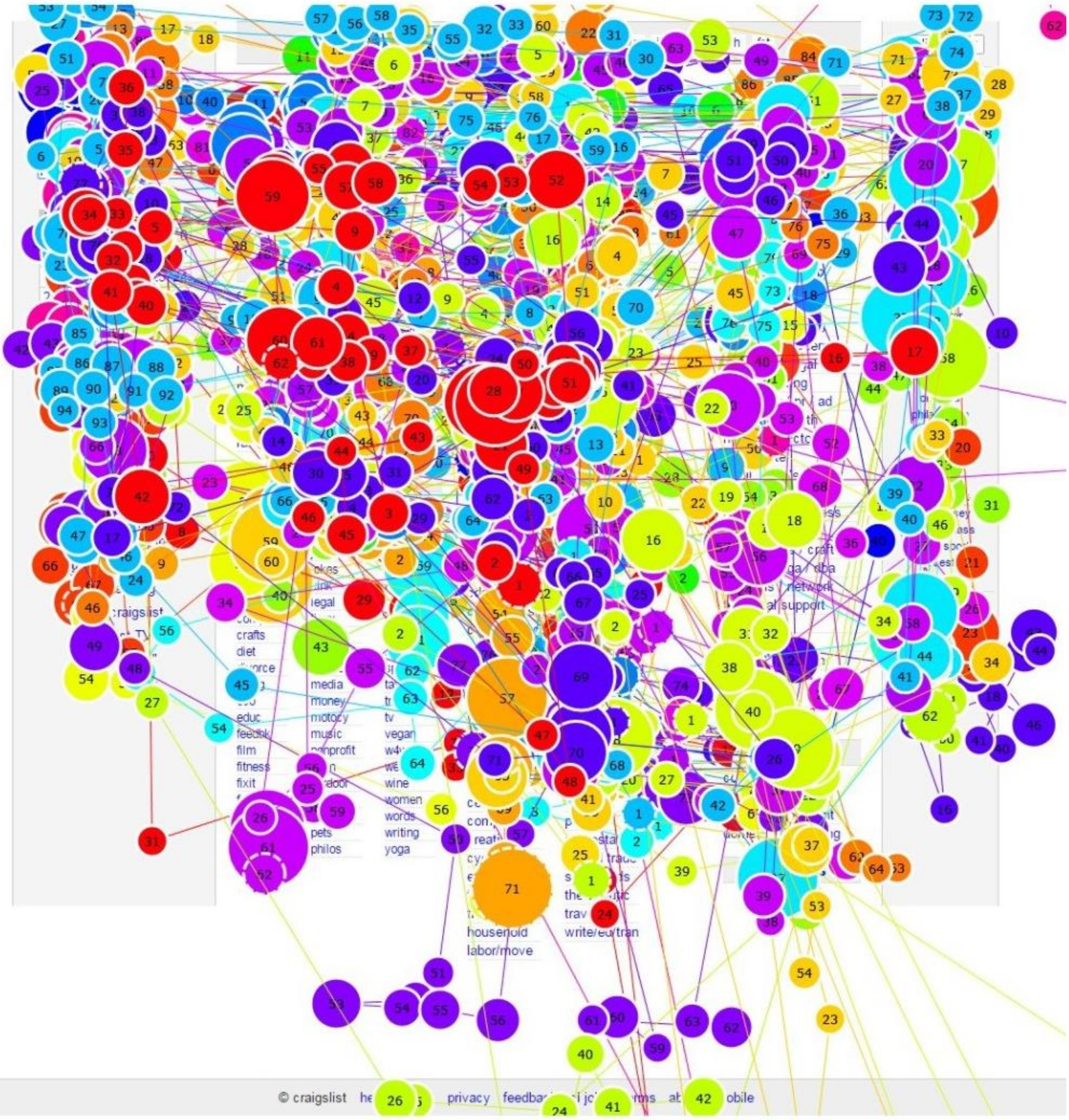

Fig. 7. The starting point of gazing for the English advertising website for all users by eye-tracking software

- Analyzing users' visual attention to decipher the interaction between CD level of a task and the advertising features reveals that users are attracted to the ad with any of the features later than the easy task in a high CD task. Further, users spend more time and view the ads more frequently in the hard task. This finding indicates that a hard task prevents users from noticing the ads sooner. However, on completion of previous research [3] and unlike the limited capacity theories, this study shows that ads are more successful in getting through the cognitive load and being noticed by users while doing the hard task than the easy task. Future research could study more complicated combinations of the features to attain higher promotion in the eye-tracking metrics.

This study reveals that advertising features and CD levels of the task have different impacts on users' visual behavior based on their gender. This research contends the following differences:

- In the simple task, females pay more frequent attention and allocate more time to the warm-colored and larger ads than males. This new finding points out the importance of size and color attribute in attracting female users' attention.

TABLE VI. LINEAR REGRESSION MODEL FOR *F3* AND *F9* IN THE LOW CD TASK

		Model	Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	t	p
TTFF	**F3**	*(Constant)*	11.63	3.06	3.79	0.00
		G	-10.78	5.86	-1.83	0.09

TABLE VII. LINEAR REGRESSION MODEL FOR *F3* IN THE HIGH CD TASK

		Model	Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	t	P
TFD	**F3**	*(Constant)*	3.42	0.63	5.39	0.00
		G	-2.62	0.89	-2.93	0.01
	F9	*(Constant)*	1.24	0.20	5.98	0.00
		G	-0.79	0.32	-2.47	0.02
FC	**F3**	*(Constant)*	11.14	2.28	4.88	0.00
		G	-7.71	3.22	-2.39	0.03
	F9	*(Constant)*	5.54	0.94	5.85	0.00
		G	-3.42	1.45	-2.34	0.03

TABLE VIII. TABLE VIII. IMPACT OF TASK CD LEVEL AND FEATURES ON USERS' VISUAL BEHAVIOR

	Source	SS	df	F	p
TTFF	**Intercept**	10989.86	1	233.89	0.00
	Fi	2133.25	9	5.04	0.00
	CDL	461.23	1	9.81	0.00
	Fi×CDL	953.20	9	2.25	0.01
TFD	**Intercept**	665.35	1	141.19	0.00
	Fi	237.31	9	5.59	0.00
	CDL	82.95	1	17.60	0.00
	Fi×CDL	126.54	9	2.98	0.00
FC	**Intercept**	7722.70	1	268.73	0.00
	Fi	851.69	9	3.29	0.00
	CDL	431.09	1	15.00	0.00
	Fi×CDL	585.36	9	2.26	0.01

TABLE IX. RESULTS OF DECISION TREE AND SVM ALGORITHMS ON USERS' GAZE PATTERN DATASET

TABLE X. *CLASSIFIERS*

		Decision stump tree		J48 decision tree		SVM		
Random seed		1	6	1	17	***SVM type***	C-SVC(classification)	nu-SVC(classification)
Cost-sensitive evaluation		Disable	Disable	Disable	Disable	***Kernel type***	Radial basis function	Polynomial
Preserve order for % Split		Disable	Disable	Disable	Disable	***Probability estimates***	FALSE	TRUE
Correctly classified instances		60%	61.66%	58.33%	61.66%	***Correctly classified instances***	63.33%	80%
Incorrectly classified instances		40%	38.33%	41.66%	38.33%	***Incorrectly classified instances***	36.66%	20%
Weighted average	**TP rate**	60%	61.70%	58.30%	61.70%	***Weighted average*** ***TP rate***	63.30%	80.00%
	FP rate	40%	38.30%	41.70%	38.30%	***FP rate***	36.70%	20.00%
	Precision	65%	78.30%	58.60%	62%	***Precision***	63.60%	80.00%
	Recall	60%	61.70%	58.30%	61.70%	***Recall***	63.30%	80.00%
	F-Measure	56%	55.10%	58%	61.40%	***F-Measure***	63.20%	80.00%

- In the hard task, coloring the words in warm colors makes it catchier at first glance for males than females.

This research intelligently matches a suitable ad for the tasks based on their CD level of them. This study claims the following classification:

- Tasks with different CD levels are classified by attaining the users' location of the first five gaze points with 80% accuracy. This finding reveals that a more efficient ad could be used for them only with users' first five gaze points data.

Considering the mother tongue language direction of users in designing an appropriate website for them, induces some challenges. This study recognizes differences among users based on this quality. Opposed with the assumed F-pattern, people with the right to left direction mother tongue language start gazing at the English advertising website mainly from the middle of the webpage. Therefore, a more suitable design of an English website can improve the user experience.

This study demonstrates the effects of many advertising features on users' two types of tasks to make them more attractive. Moreover, it intelligently adjusts features of ads with types of tasks. Also, it figures out differences in gazing patterns among users for one of the most popular advertising websites in the USA, with the purpose of promotion in the design of English websites in the future. However, further research on specific conditions of the studied ads is essential. Hopefully, this current research affects future scholarship in this field.

References

[1] AdNews (2013, June13). No one's paying attention to digital ads…well, 8% are Retrieved from http://www.adnews.com.au/adnews/no-one-s-paying-attention-to-digital-ads-well-8-are.

[2] Bang, Hyejin, and Bartosz W. Wojdynski. "Tracking users' visual attention and responses to personalized advertising based on task cognitive demand." Computers in Human Behavior 55 (2016): 867-876.

[3] Dehghani, Milad, et al. "Evaluating the influence of YouTube advertising for attraction of young customers." Computers in human behavior 59 (2016): 165-172.

[4] Kumar, Naveen, Varun Maheshwari, and Jyoti Kumar. "A comparative study of user experience in online social media branding web pages using eye tracker." 2016 international conference on advances in human machine interaction (HMI). IEEE, 2016.

[5] Buscher, Georg, Edward Cutrell, and Meredith Ringel Morris. "What do you see when you're surfing?: using eye tracking to predict salient regions of web pages." Proceedings of the SIGCHI conference on human factors in computing systems. ACM, 2009.

[6] Sørum, Hanne. "Design of Public Sector Websites: Findings from an Eye Tracking Study Emphasizing Visual Attention and Usability Metrics." International Conference on Electronic Government and the Information Systems Perspective. Springer, Cham, 2016.

[7] Zhang, Jin-Ting, et al. "A new test for functional one-way ANOVA with applications to ischemic heart screening." Computational Statistics & Data Analysis 132 (2019): 3-17

[8] Ponzio, Francesco, et al. "A human-computer interface based on the "voluntary" pupil accommodative response." International Journal of Human-Computer Studies 126 (2019): 53-63.

[9] Kaur, Akriti, Ashutosh Agrawal, and Pradeep Yammiyavar. "Exploring 3D Interactions for Number Entry and Menu Selection in Virtual Reality Environment." Research into Design for a Connected World. Springer, Singapore, 2019. 781-791.

[10] Jerome, Theresa, Leong Wai Shan, and Kok Wei Khong. "Online advertising: a study on Malaysian consumers." Available at SSRN 1644802 (2010).

[11] Brajnik, Giorgio, and Silvia Gabrielli. "A review of online advertising effects on the user experience." International Journal of Human-Computer Interaction 26.10 (2010): 971-997.

[12] Rodgers, Shelly, and Esther Thorson. "The interactive advertising model: How users perceive and process online ads." Journal of interactive advertising 1.1 (2000): 41-60.

[13] Faber, Ronald J., Mira Lee, and Xiaoli Nan. "Advertising and the consumer information environment online." American Behavioral Scientist 48.4 (2004): 447-466.

[14] Rodgers, Shelly, and Esther Thorson. "The interactive advertising model: How users perceive and process online ads." Journal of interactive advertising 1.1 (2000): 41-60.

[15] ACM SIGCHI: The Association for Computing Machinery Special Interest Group on Computer-Human Interaction (1992). www.sigchi.org. Accessed 14 Mar 2016

[16] Frøkjær, Erik, and Kasper Hornbæk. "Metaphors of human thinking for usability inspection and design." ACM Transactions on Computer-Human Interaction (TOCHI) 14.4 (2008): 20.

[17] International Organization for Standardization (ISO) 9241-11: Ergonomic Requirements for Office Work with Visual Display Terminals (VDTs) – Part 11: Guidance on Usability (1998)

[18] Berger, Jonah, and Katherine L. Milkman. "What makes online content viral?" Journal of marketing research 49.2 (2012): 192-205.

[19] Yu, Jay, and Brenda Cude. "'Hello, Mrs. Sarah Jones! We recommend this product!'Consumers' perceptions about personalized advertising: comparisons across advertisements delivered via three different types of media." International journal of consumer studies 33.4 (2009): 503-514.

[20] Lang, A. (2006). Using the limited capacity model of motivated mediated message processing to design effective cancer communication messages. Journal of Communication,56(1),57–S80.

[21] Petty, R.E., Barden, J., & Wheeler, S.C.(2002). The elaboration likelihood model of persuasion: health promotions that yield sustained behavioral change. InR.J. DiClemente, R.A. Crosby, &Kegler (Eds.), Emerging theories in health promotion practice and research(pp.71–99). San Francisco: Jossey-Bass.

[22] Malheiros, Miguel, et al. "Too close for comfort: A study of the effectiveness and acceptability of rich-media personalized advertising." Proceedings of the SIGCHI conference on human factors in computing systems. ACM, 2012.

[23] Tam, Kar Yan, and Shuk Ying Ho. "Web personalization as a persuasion strategy: An elaboration likelihood model perspective." Information systems research 16.3 (2005): 271-291.

[24] Kreuter, Matthew W., and Ricardo J. Wray. "Tailored and targeted health communication: strategies for enhancing information relevance." American journal of health behavior 27.1 (2003): S227-S232.

[25] Pavlou, Paul A., and David W. Stewart. "Measuring the effects and effectiveness of interactive advertising: A research agenda." Journal of Interactive Advertising 1.1 (2000): 61-77.

[26] Baek, Tae Hyun, and Mariko Morimoto. "Stay away from me." Journal of advertising 41.1 (2012): 59-76.

[27] Howard, Daniel J., and Roger A. Kerin. "The effects of personalized product recommendations on advertisement response rates: The "try this. It works!" technique." Journal of consumer psychology 14.3 (2004): 271-279.

[28] Li, Cong. "When does web-based personalization really work? The distinction between actual personalization and perceived personalization." Computers in Human Behavior 54 (2016): 25-33.

[29] Phelps, Joseph E., Giles D'Souza, and Glen J. Nowak. "Antecedents and consequences of consumer privacy concerns: An empirical investigation." Journal of Interactive Marketing 15.4 (2001): 2-17.

[30] Sacirbey, O. "Privacy Concerns, Advertising Hamper E-commerce." IPO Reporter 24.35 (2000): 11-11.

[31] Van Doorn, Jenny, and Janny C. Hoekstra. "Customization of online advertising: The role of intrusiveness." Marketing Letters 24.4 (2013): 339-351.

[32] Kempf, DeAnna S., Kay M. Palan, and Russell N. Laczniak. "Gender differences in information processing confidence in an advertising context: A preliminary study." ACR North American Advances (1997).

[33] Shavitt, Sharon, Pamela Lowrey, and James Haefner. "Public attitudes toward advertising: More favorable than you might think." Journal of advertising research 38.4 (1998): 7-22.

[34] Sweller, John. "Cognitive load during problem solving: Effects on learning." Cognitive science 12.2 (1988): 257-285.

[35] Sweller, John, Jeroen JG Van Merrienboer, and Fred GWC Paas. "Cognitive architecture and instructional design." Educational psychology review 10.3 (1998): 251-296.

[36] Rodgers, Shelly, and Esther Thorson. "The interactive advertising model: How users perceive and process online ads." Journal of interactive advertising 1.1 (2000): 41-60.

[37] Hoffman, Donna L., and Thomas P. Novak. "Marketing in hypermedia computer-mediated environments: Conceptual foundations." Journal of marketing 60.3 (1996): 50-68.

[38] Wise, Kevin, Hyo Jung Kim, and Jeesum Kim. "The effect of searching versus surfing on cognitive and emotional responses to online news." Journal of Media Psychology 21.2 (2009): 49-59.

[39] Wolford, George, and Fred Morrison. "Processing of unattended visual information." Memory & Cognition 8.6 (1980): 521-527.

[40] Nielsen, Jakob. " F-Shaped Pattern For Reading Web Content, Jakob Nielsen's Alertbox." http://www. useit. com/alertbox/reading_pattern. html (2006).

[41] Nielsen, J.: Designing Web Usability. New Riders Publishing, Indianapolis (2000). ISBN 1-56205-810-X

[42] Johansen, Sune Alstrup, and John Paulin Hansen. "Do we need eye trackers to tell where people look?." CHI'06 Extended Abstracts on Human Factors in Computing Systems. ACM, 2006.

[43] Ehmke, Claudia, and Stephanie Wilson. "Identifying web usability problems from eye-tracking data." Proceedings of the 21st British HCI Group Annual Conference on People and Computers: HCI... but not as we know it-Volume 1. British Computer Society, 2007.

[44] Bojko, Aga. Eye tracking the user experience: A practical guide to research. Rosenfeld Media, 2013, P.27.

[45] Byrne, Michael D., et al. "Eye tracking the visual search of click-down menus." Proceedings of the SIGCHI conference on Human Factors in Computing Systems. ACM, 1999.

[46] Jacob, Robert JK, and Keith S. Karn. "Eye tracking in human-computer interaction and usability research: Ready to deliver the promises." The mind's eye. North-Holland, 2003. 573-605.

[47] Carpenter, P. A., and M. A. Just. "Linguistic influences on picture scanning." Eye movements and psychological processes (1976): 459-472.

[48] Graf, W., and H. Krueger. "Ergonomic evaluation of user-interfaces by means of eye-movement data." Proceedings of the third international conference on human-computer interaction, Vol. 1 on Work with computers: organizational, management, stress and health aspects. Elsevier Science Inc., 1989.

[49] Wedel, Michel, and Rik Pieters. "A review of eye-tracking research in marketing." Review of marketing research. Routledge, 2017. 123-147.

[50] Kong, Shaojun, et al. "Web advertisement effectiveness evaluation: Attention and memory." Journal of Vacation Marketing 25.1 (2019): 130-146.

[51] Muñoz-Leiva, Francisco, Janet Hernández-Méndez, and Diego Gómez-Carmona. "Measuring advertising effectiveness in Travel 2.0 websites through eye-tracking technology." Physiology & behavior 200 (2019): 83-95.

[52] Okano, Masao, and Masami Asakawa. "Eye tracking analysis of consumer's attention to the product message of web advertisements and TV commercials." 2017 5th International Conference on Cyber and IT Service Management (CITSM). IEEE, 2017.

[53] Weizhe, Li, and Fei Hu. "Eye-Tracking assisted Human-Computer Interaction on E-Shopping Website: A Preliminary Study." (2020).

[54] Jagadale, Pooja Ganesh. "Role of Eye Tracking System To Enhance Life Of Disable People." International Research Journal of Modernization in Engineering Technology and Science 2(11) (2020):715-719.

[55] Agarkhed, Jayashree, et al. "Human Computer Interaction System Using Eye-Tracking Features." 2020 IEEE Bangalore Humanitarian Technology Conference (B-HTC). IEEE, 2020.

[56] Brosens, Jacques, Funmi Adebesin, and Rendani Kruger. "In the eye of the beholder: Teaching user-centered design to information and communication technology students with the help of eye tracking." Handbook of research on diverse teaching strategies for the technology-rich classroom. IGI Global, 2020. 296-318.

[57] Gwizdka, Jacek. "Distribution of cognitive load in web search." Journal of the American Society for Information Science and Technology 61.11 (2010): 2167-2187.

Yasaman Mashhadi Hashem Marandi is a graduate student with M.Sc in Computer science from University of Tehran (2015- 2018) and B.Sc in Computer science from Kharazmi University (2010-2014). Her works focus specifically on Human-computer interaction. She also has an interest in logic.

Hedieh Sajedi is an Associate Prof. at Department of Computer Science, School of Mathematics, Statistics and Computer Science, College of Science, University of Tehran.