Information Fusion 38 (2017) 12–21

Contents lists available at ScienceDirect

Information Fusion

journal homepage: www.elsevier.com/locate/inffus

Full Length Article

Biometric information fusion for web user navigation and preferences analysis: An overview

Gino Slanzi, Gaspar Pizarro, Juan D. Velásquez*

Department of Industrial Engineering, Universidad de Chile, República 701 P.O. Box: 8370439, Santiago, Chile

ARTICLE INFO

Article history:
Received 26 September 2016
Revised 4 February 2017
Accepted 6 February 2017
Available online 11 February 2017

Keywords:
Web usage mining
Information fusion
Biometric data

ABSTRACT

Throughout the years having knowledge of Web users' interests, navigational actions and preferences has gained importance due to the objectives of organizations and companies. Traditionally this field has been studied from the Web Mining perspective, particularly through the Web Usage Mining (WUM) concept, which consists of the application of machine learning techniques over data originated in the Web (Web data) for automatic extraction of behavioral patterns from Web users. WUM makes use of data sources that approximate users' behavior, such as weblogs or clickstreams among others; however these sources imply a considerable degree of subjectivity to interpret. For that reason, the application of biometric tools with the possibility of measuring actual responses to the stimuli presented via websites has become of interest in this field. Instead of doing separate analyses, information fusion (IF) tries to improve results by developing efficient methods for transforming information from different sources into a single representation, which then could be used to guide biometric data fusion to complement the traditional WUM studies and obtain better results. This paper presents a survey of Biometric IF applied to the WUM field, by first defining WUM and its main applications, later explaining how the Biometric IF could be applied and finally reviewing several studies that apply this concept to WUM.

1. Introduction

Throughout the last two decades, the Web has grown quickly in usage and penetration, which has led to changes in the behavior of people, businesses and organizations. It now seems almost natural to use the Web daily to look for any sort of information, to get products and services and to socially interact through different platforms and websites. Hence, companies and organizations make their best efforts to obtain a position within this network in order to attract and retain users and customers. To achieve this objective it is necessary to be interesting and effective on the Web by having more engaging websites than competitors.

Nevertheless, the task of having outstanding websites is not trivial at all. On the contrary, having specific knowledge about the needs of potential users and customers along with the ability to establish personalized services to satisfy these needs is essential [1]. This is directly related to how people interact with websites, how they behave while browsing, what their preferences are and what zones drive their attention, ideas encompassed by the concept of Web Usage Mining (WUM).

Web Usage Mining, defined as "*the application of machine learning techniques over Web data for automatic extraction of behavioural patterns from Web users*" [2], makes use of different techniques for modeling Web user experience, exploiting diverse Web data sources like Web server logs, clickstreams, and user profiles, among others [4]. WUM developments aim to produce behavioral patterns that can be used as inputs to applications such as recommendation systems, visualization, Web analytics or report generation tools [6].

Typical WUM sources of data have led these studies in the direction of a considerable level of subjectivity. There is a need for more accurate and objective data to describe the navigation and preferences of Web users. As a result, biometric tools like eye trackers (ET) or electroencephalography (EEG) [40,62], have gained presence in this field, because they allow researchers to measure neurophysiological responses to the stimuli presented via websites and thus, try to understand what is actually happening when users browse Web pages.

There is another research approach focused on the analysis of keystroke dynamics. Several studies have been performed with the objective of describing and classifying keystroke biometrics for identity verification [82,83,85,87,88]. Although these techniques have not been highly used in Web navigation and preferences

* Corresponding author.
E-mail addresses: gslanzi@ing.uchile.cl (G. Slanzi), gaspar.pizarro.v@wic.uchile.cl (G. Pizarro), jvelasqu@dii.uchile.cl (J.D. Velásquez).

http://dx.doi.org/10.1016/j.inffus.2017.02.006
1566-2535/

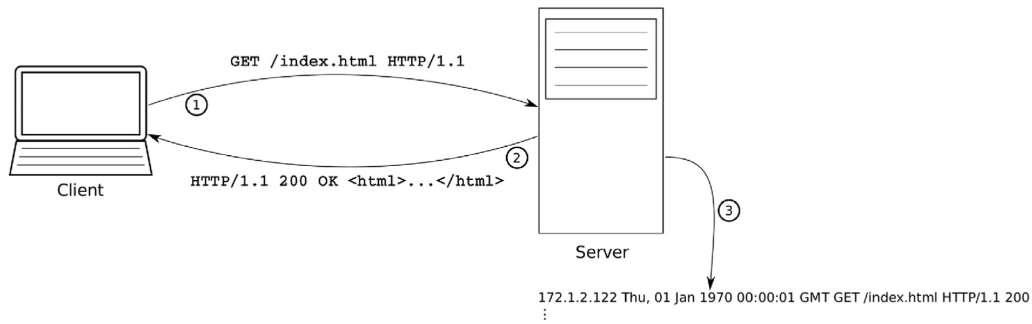


Fig. 1. HTTP protocol simplified diagram. First, a client sends a request with method and path, then the server sends a response with the status code and the requested content. Finally the server adds an entry to its log, with (at least) a timestamp, the client's IP, the HTTP header and the response's status code.

analysis and lay beyond the scope of this article, it is interesting to mention their existence, for information purposes.

On the other hand, information fusion (IF) is defined as "*the study of efficient methods for automatically or semi-automatically transforming information from different sources and different points in time into a representation that provides effective support for human or automated decision making*" [8,21–24]. IF first applications and methods were established in the military field, however, over time several non-military uses were found, for example, in the industrial and commercial sectors [7]. In this regard, this paper intends to review the biometric data analysis approach for WUM, addressing it as an existing application of IF.

The rest of the paper is structured as follows. In Section 2 the traditional WUM methodology and principal applications are described. Next, in Section 3 the biometric information fusion as a complement for WUM is depicted, presenting a review of the state of the art on this topic and proposing a conceptual framework for guiding the biometric fusion process in the WUM field. Finally, in Section 4 we examine some literature reviews related to this work.

2. Web user navigation and preferences analysis in general

In this section we describe the traditional Web Usage Mining methodology by briefly depicting each step from a Data Mining perspective and discussing some emblematic practical applications.

2.1. Web mining overview

Web Mining is the area of Data Mining which deals with the extraction of interesting knowledge from the World Wide Web [9]. The diverse data sources that originate in this domain lead to the possibility of differentiating Web Mining into three main areas of study [10,13]. On the one hand, there is Web Content Mining (WCM), related to the mining of knowledge from content present on Web pages such as text, images, videos, etc. [11]. Furthermore, Web Structure Mining (WSM) focuses on the hyperlink structure of the Web [12]. Finally, Web Usage Mining Web (WUM) refers to the automatic discovery and analysis of patterns in clickstream and associated data collected or generated as a result of user interactions with Web resources on one or more Web sites [6,66]. In this work we are centered in the analysis of Web users' navigation and preferences, which is directly related with WUM. Fig. 1 shows how the HTTP protocol works and how the data for WUM originates.

2.2. Web usage mining

The WUM process could be described as a typical Data Mining problem, using Web content as input. This means that a typical WUM process can be divided into four stages, that are described in the next sections.

2.2.1. Data collection

This step considers the selection of the right data sources to accomplish the objectives of the study. In the specific case of WUM, the classical data sources are Web servers, clients that connect with servers, or intermediary sources like proxy servers.

Web logs correspond to files in which the servers record every user request for resources of a particular website. These archives contain information about users' navigational behavior, including IP, a timestamp, resource, server, browser type and version, request, method and status, among other data [15,16]. A log file is shown in Fig. 2.

2.2.2. Preprocessing

In this stage, the collected data is processed in order to remove the cases that are not useful for the purposes of the study. Depending on the source, the data would be preprocessed differently. Typically, WUM preprocessing includes several sub-steps such as:

Data filtering This is the process of cleaning the data in order to remove entries not referred to users, such as static files like graphics or sound files, or automatized crawlers, like Google or Bing's crawlers.

User identification According to [16], this is the process for identifying individual users by observing their IP address. A basic heuristic to do that is to use the combination of IP and User-Agent fields in the web log to identify a single user [20].

Pageview identification Page-view identification is the process of determining which page file accesses contribute to a single browser display [3].

Sessionization This is the process of segmenting the user activity record of each user activity record of each user into sessions, each one representing a single visit to the site [6].

Path completion This is the process of determining if there are important accesses that are not recorded in the web log [19]. This can happen because of the local cache that is masking user requests or because of the user of the "back" button.

These techniques have been surveyed previously in [17–20]. At the end of this phase, according to [6], we have a list of page views and user transactions:

$$t = [(p_1, w(p_1)), (p_2, w(p_2)), \ldots, (p_n, w(p_n))] \quad (1)$$

where t is a *transaction*, p_i is a *pageview* and $w(p_i)$ is a weight associated to p_i, for instance, the time a user spent with that page view. This type of data can be used as input for a posterior data mining step, with some other preprocessing. For example, according to [6], if we index the n page views of a site with an index i, we

#	IP	Time	Method/URL/Protocol	Status	Bytes	Referrer	User-Agent
1	165.182.168.101	1415588443	GET /robots.txt HTTP/1.0	200	173	-	msnbot/1.0 (+http://search.msn.com/msnbot.htm)
2	165.182.168.101	1415588443	GET /home HTTP/1.0	200	14199	-	msnbot/1.0 (+http://search.msn.com/msnbot.htm)
3	204.231.180.195	1415588460	GET /p1.aspx HTTP/1.1	200	3171	p3.htm	Mozilla/4.0 (compatible; MSIE 5.5; Windows 98; SAFEXPLORER TL)
4	204.231.180.123	1415588462	GET /p2.htm HTTP/1.1	200	8090	p1.htm	Mozilla/4.0 (compatible; MSIE 6.0; Windows NT 5.1; SV1)

Fig. 2. Typical web log. For the sake of space, some fields are omitted.

can make a vector $\vec{v} = (w(p_i))_{i\in\{1,\ldots,n\}}$. This vector can be passed to a domain-independent (not necessarily associated to a web mining task) data mining model or algorithm.

2.2.3. Pattern discovery

The third phase considers the utilization of different methods and algorithms for pattern recognition and extraction. Depending on the objectives, statistical, machine learning or data mining methodologies could be chosen in this step. The most common techniques applied in WUM include clustering (segmentation), classification (prediction), associative and sequential (navigational) analyses.

2.2.4. Pattern and results analysis

As a final part of the process, analyzing the discovered patterns is a crucial step, since decisions are made out of these results. The WUM process could be used to generate different outcomes, such as personalization, recommender systems and content and structure improvements, among others.

2.3. Web usage mining practical applications

According to [2], there are three main areas in which Web Usage Mining plays a key role. These areas correspond to Adaptive Websites, Web Personalization, and Recommendation, the first one being the basis of the WUM applications. As the pretensions of this work are not to provide a thorough review of these, we will include a succinct description of each mentioned area.

2.3.1. Adaptive websites

Literally, an Adaptive Website (AWS) is a site that has the ability of adapting in regard to the needs or the preferences of its users. As stated in [25], AWS are "*sites that semi-automatically improve their organization and presentation by learning from visitor access patterns*". This kind of website has the capability of modifying its structure and content based on the individual user behavior [26]. These changes may alter the website or run the risk of user rejection, in which case it is advisable to implement them as recommendations for the visitor or the expert who maintains the website [27].

2.3.2. Web personalization

Nowadays, organizations face the necessity of attracting and retaining Web users as eventual customers. A plausible way to deal with this is through Web Personalization, which can be defined as an action that tailors the Web experience to a particular user or set of users. Actions can range from making the presentation more pleasing to predicting users' needs and providing customized information [28]. The site is personalized through the highlighting of existing hyperlinks, the dynamic insertion of new hyperlinks that seem to be of interest for the current user, or even the creation of new index pages [29]. The application of personalization is surveyed in [14].

The objective of the personalization process is to provide users with whatever they need at any time, without requiring them to ask for it explicitly. That is to say, inferring the actual requirements based on either previous or current navigational behavior or interactions [30].

2.3.3. Recommendation

Recommendation systems can be generally defined as software tools and techniques with the ability to provide suggestions of items that could be of use to a user, or in other words, to help users select an item from a large space of possible options [77–79]. Traditionally, these systems take Web users' navigational or rating data as input to data mining or machine learning algorithms to find patterns that represent aggregate user models. Then, whenever a new user enters the system, their behavior will automatically be matched with those patterns, and interesting items selected as recommendations [80].

3. Biometric information fusion for web user navigation and preferences analysis

In this section we describe how the traditional WUM approach has been enriched through the incorporation of biometric information fusion at different levels. We briefly overview the information fusion field, then we describe some of the most-used technologies in this regard. Afterwards, we depict how biometric information is fused to enrich traditional WUM approaches, focusing on the broad levels of fusion of data sources and fusion of resources and techniques. Finally, we propose a conceptual framework with the objective of guiding the biometric information fusion process in this specific scenario.

3.1. Information fusion overview

The literature presents several definitions that could be associated with Information fusion. All of them coincide in terms of using concepts such as fusing, merging and aggregating or integrating information provided by multiple sources in order to generate better answers than using those sources separately.

In 1991, the Joint Directors of Laboratories (JDL) published a data fusion lexicon in which data fusion was defined as "*A process dealing with the association, correlation, and combination of data and information from single and multiple sources to achieve refined position and identity estimates, and complete and timely assessments of situations and threats as well as their significance*" [73]. Hall and Llinas in 1997 manifest that "*Data fusion techniques combine data from multiple sensors, and related information from associated databases, to achieve improved accuracies and more specific inferences than could be achieved by the use of a single sensor alone*" [32]. Similarly, based on his previous work, Wald adopted the following definition: "*data fusion is a formal framework in which are expressed means and tools for the alliance of data originating from different sources. It aims at obtaining information of greater quality; the exact definition of 'greater quality' will depend upon the application*" [33].

More recently, Boström et al. [8], in attempting to establish a formal definition of IF as a field of research, reviewed more than 30 previous works regarding this topic, and came out with the next statement for defining it: "*Information fusion is the study of efficient methods for automatically or semi-automatically transforming information from different sources and different points in time into a representation that provides effective support for human or automated decision making*".

Khaleghi et al. performed a thorough review of the state of the art in multi-sensor data fusion. They addressed its definition, ben-

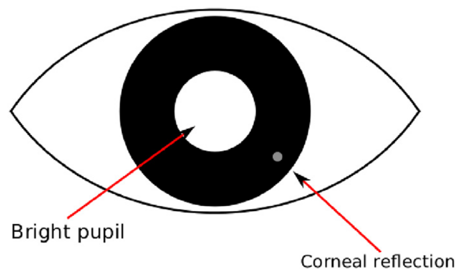


Fig. 3. Features used by an infrared-based eye tracker. Based on [54].

efits, challenges and algorithms, as well as the emerging fusion paradigms and the most recent and ongoing research [34].

It is important to remember that the first information fusion applications and methods were in the military realm. Afterwards, other sorts of non-military uses were absorbed into the IF point of view, including within the industrial and commercial sectors [7]. Based on this last fact, we intend to formalize and review the application of biometric information fusion to enrich traditional WUM analysis.

3.2. Diverse technologies to choose from

In this study we are focused on reviewing studies that tend to fuse diverse biometric data in order to complement the traditional WUM approach. These studies commonly make use of technologies capable of recording eye movements and brain activation. The most-used technologies correspond to the following:

3.2.1. Eye tracking

This is a technique whereby an individual's eye movements are measured to determine where a person is looking in real time and the sequence in which their eyes are shifting from one location to another [50].

The Eye-Tracker system available today measures point-of-regard by the corneal-reflection and pupil-center method. The system consists of a standard desktop computer with an infrared camera mounted next to a display monitor, with an image processing software to identify and track eye features. Infrared light from an LED embedded in the infrared camera is first directed at the retina, creating reflections that allow the eyes to be tracked [54]. Fig. 3 shows the features captured by an eye-tracker.

As said before the pupil-center method needs a source of infrared light and a camera. These devices can be deployed attached to a monitor, and the user has to be in a fixed position for it to work, or they can be deployed in a pair of glasses, offering more freedom for the user. For instance in [51], they offer devices with both ways of deployment.

Regarding low-level data, the measurements given by this system are the position and size of the pupil. For instance, the sampling rate of the eye tracker we dispose is approximately 20 Hz, whereas the sampling rate of the one of [51] can be of 50 or 100 Hz, depending on the model. This signal can have noise from blinking, where the eye tracker loses the features for capturing the eye gaze, and "saccades", which are quick eye movements occurring between more stationary points, called fixations, which is the data we have after cleaning. The data from the fixations can be aggregated, in one eye-tracking session or with several ones, in order to determine the most watched points in an image, which makes a fixation map [86]. These two types of data are shown in Fig. 4.

There are different studies where researchers have used such techniques as eye tracking to analyze Web user behavior on specifics tasks, in order to improve website content and design. For example, in 2007, Cutrell and Guan researched how to improve the Web search services using eye-tracking tools, concluding that these are improved when information is added to contextual snippets, but performance is degraded for gravitational tasks [56].

3.2.2. Electroencephalography (EEG)

This is a medical-imaging technique that reads scalp electrical activity generated by brain structures. The EEG is measured directly from the cortical surface and mostly involves the current that flows during synaptic excitations of the dendrites of many pyramidal neurons in the cerebral cortex. By this means, EEG allows a measurement of the changes in the electrical potential over time between a signal electrode and a reference electrode [57,58].

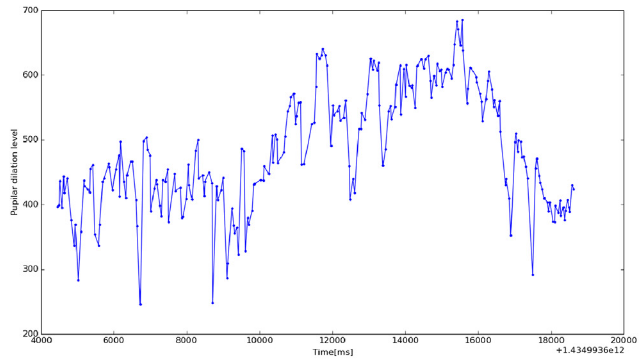
An EEG system consists of electrodes, amplifiers and a recording device. There are systems with 14, 32, 64 or 128 electrodes, distributed in a standard localization defined by the International Federation of Clinical Neurophysiology, which developed the 10–20 electrode placement system [84]. Electrodes commonly consist of silver or silver chloride disks, 1 to 3 mm in diameter, with long flexible leads that can be plugged into an amplifier. This amplifier is connected to a recording device, generally a computer, which shows the signals as a wave function over time. Two examples of this system are EMOTIV's Epoc+, which measures 14 channels at 128 or 256 samples per second [53], and Biosemi's ActiveTwo system, which is a generic biopotential measurement system (not only for EEG, as the Epoc+) that can measure up to 280 channels at 8192 samples per second [52]. Fig. 5 shows data obtained from an EEG device. A standard for EEG data (and other biosignals) is the European Data Format (EDF) [89], which, roughly speaking, specifies details of a generic signal, such as extreme values, dimensions, start and end times, number of samples (so the sampling rate can be derived) and some annotations.

The main uses of the EEG are in clinical settings to detect gross pathologies and epilepsies and in research facilities. But now, the newest research that uses this technique is applied to study human behavior in other disciplines. In 2006, Liang and colleagues developed a machine-learning algorithm referred to as Extreme Learning Machine (ELM) that does classifications of five mental tasks from subjects using EEG signals. ELM was compared with such other algorithms as Support Vector Machines (SVMs) and Back Propagation Neural Network (BPNN). Their results showed that the classification accuracy of ELM is similar to SVM and BPMN. ELM also took less time to determine the parameters of the classifier than other method [59].

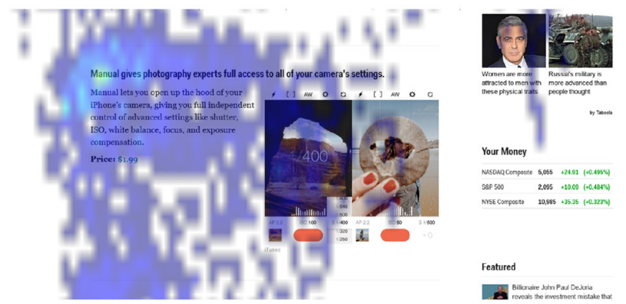
3.2.3. Functional magnetic resonance image (fMRI)

This is a non-invasive technology for imaging the activation of brain areas by various types of physical sensation or activity, problem solving and movement. fMRI can map the brain, through detection of changes in levels of blood oxygen in brain areas, generated by brain activities [60]. The change is produced by the oxygen saturation of the hemoglobin, denominated the blood oxygenation level-dependent (BOLD) effect, which is described with mathematical models that explain how its signal is in function of O_2 levels [81]. The level of oxygenation influences the variation in the blood's magnetic resonance (MR) signal. When the concentration of deoxyhaemoglobin (hemoglobin without oxygen) is low, the intensity of the images captured by fMRI machines increases, and vice versa. The data obtained by this method are 3-dimensional images of the brain and its neurons activation levels. Compared to the technologies described in Sections 3.2.1 and 3.2.2, this technique is the most massive on data of the three. For instance, the Haxby dataset [5] has, for every subject, approximately 300 MB of data, corresponding to only 30 s of MRI measurements. Fig. 6 shows a frame from the Haxby dataset.

In 2008, Ghebreab et al. [61] developed a methodology that allows recognition of behavior patterns, using fMRI, to predict dif-



(a) Pupil dilation recorded from an eye-tracker during an experiment, where blinks and saccades were removed. Source: own elaboration.



(b) Fixation map on section of a webpage. Source: own elaboration.

Fig. 4. Eye-tracker-originated data.

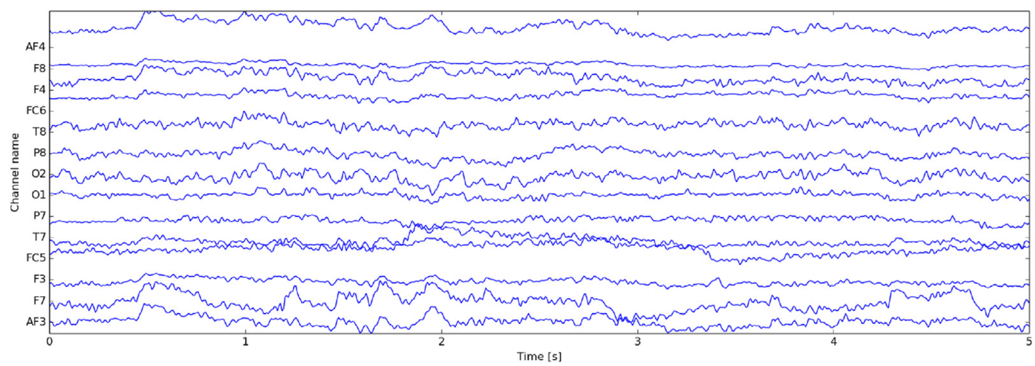


Fig. 5. 14-channel signal from an EMOTIV device. Source: own elaboration.

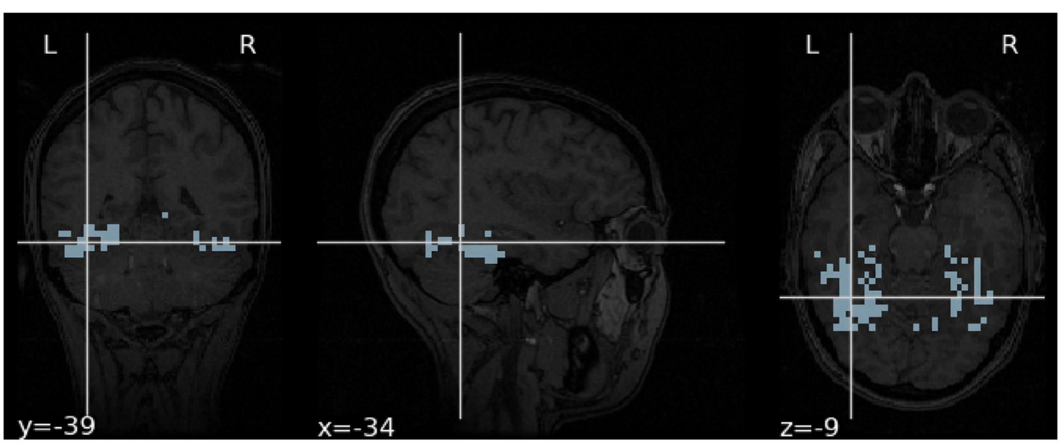


Fig. 6. Data originated from an fMRI machine, from [5]. The gray pixels show the regions of the brain activated in a certain time, with an anatomical image as background. Source: own elaboration.

ferent human brain states through the study of complex stimuli. This method identifies important covariance between spatially-distributed brain responses (measured with fMRI technology) and stimuli.

3.3. Fusion of data sources

Several studies combine different data sources to complement traditional WUM research or experiments from another perspective. Accordingly, this subsection includes studies that fuse, for example, eye-tracking ocular positioning with Web object mapping to analyze saliency.

In the work presented in [36], the authors implement an eye-tracking-based analysis in which 20 users were shown 361 web-pages while performing information foraging and inspection tasks. The main assumption was that gaze data could represent a proxy of attention. Taking that into account, they developed an analysis framework, by first generating a tool that characterizes DOM elements and then maps the users' gazes to them, and second by using extracted web features in a machine-learning setting to predict the salient elements of a web page. They also introduce the concept of mean fixation impact, which allows the identification of the set of elements that are under the gaze of the user at a certain time.

Loyola and Velásquez [38] proposed a graph-based analysis framework to study the dynamics of visual gaze from web users. Their approach consisted of modeling Web objects as nodes, and visual transitions (obtained with an eye tracker) as edges. Their results suggest that a graph-based analysis can capture, in a reliable way, the dynamics of user behavior and the identification of salient objects within a website.

In 2010, Lee and Seo performed a usability study in which typical techniques were mixed with biosignal analysis. Specifically, they compared the results of a standard usability test using user-based evaluations (user performance measurements, keystroke analysis, satisfaction questionnaires and interviews), with the results obtained from analyzing EEG and ECG data. They found that using this new biosignal-based approach was a reasonable and valuable method for web evaluation, since they obtained 70% precision in the comparison [39].

In [42] a study about how people interact with Web search engine result pages is presented by using eye tracking in combination with a systematic variation of types of tasks, quality of ads, and sequences in which ads of different quality were presented. Authors measured the effects of these variables on the distribution of visual attention and on task performance, finding significant effects for each variable.

In [47] the authors proposed a biometric-based approach for predicting Web user click intention, using pupil dilation data generated by an eye-tracking device in combination with Web object mapping as input. They intended to determine if this variable was useful to differentiate choice and not-choice states, understanding choice as a click. They also tried to generate a classification model based on pupil dilation and ocular positioning features. Their results suggest that this variable could be used from a Web Intelligence point of view as a proxy of Web user behavior, in order to generate an online recommender to improve Web site structure and content.

The work of [48] focuses on the problem of detection of users' areas of interest within Web pages. They claim that the attention level while browsing is a more reliable indication of the level of interest. Authors implemented an experiment where the EEG signals were captured in real time, while a background script in a Web browser captured the section of the web page currently being viewed, calculating the percentage of the page that the user had scrolled to. They were able to determine and map the average attention level within pages for a range of websites and users, which could be useful in diverse applications such as providing inputs of user behavior to web developers for better web design, ranking different websites or videos as per user interest, inserting ads in the regions of a web page where the user is more likely to pay attention to.

In [56] the initial results of a project regarding the feasibility of inferring web page relevance from eye-tracking data are shown. Authors perform analyses of variance as well as classification algorithms in order to predict user-perceived Web page relevance, demonstrating that it is feasible to infer document relevance from eye-tracking data on Web pages. The results indicate that the combination of eye fixation duration, pupil size and the probability of continuing reading are good predictors of Web page relevance.

In [63], the authors explored the differences between searching and surfing on cognitive and emotional responses to online news. They conducted an experiment where 92 subjects read three unpleasant news stories from a website. Half of the participants acquired their stories by searching, while the other half acquired their stories by surfing, thus, with no previous information need in mind. Heart rate, skin conductance, and corrugation activation were collected as measures of resource allocation, motivational activation, and unpleasantness, respectively, while participants read each story. Self-report valence and recognition accuracy were also measured. They found that the stories acquired by searching elicited greater heart rate acceleration, skin conductance level, and corrugation activation during reading. These stories were rated as more unpleasant, and their details were recognized more accurately than similar stories that were acquired by surfing. They also discussed the implications of these results for understanding how people process online media.

In [64] the authors developed a methodology for identifying *Website Keywords*, defined as "*a word or possibly a set of words that is used by visitors in their search process and characterizes the content of a given Web page or website*", by analyzing plain text. Afterwards this methodology was extended to include the analysis of other kinds of content present on websites, such as images or videos, and *Website Objects* which were defined as "*any structured group of words or multimedia resource within a Web page that has metadata to describe its content*". They finally defined the *Website Keyobjects* as "*a Web object or group of Web objects that attract Web users' attention and that characterize the content of a given web page or website*", which objects would be the most probable elements to be chosen or clicked on the website [65,67]. To further improve the methodology and make it more objective, Velasquez et al. incorporated an eye-tracking-based analysis for estimating the time spent on each object by each user, thus obtaining better results [68]. Finally, a pupil dilation approach was used to complement the previous work, finding that the inclusion of pupillary activity [55], although not conclusively, allows the extraction of a more robust Web Object classification, achieving a 14% increment in the overall accuracy [69].

Finally, in [72], the authors proposed a biological-based feature comparison for identifying salient Web objects. Several features extracted from eye tracking and EEG data were compared to a baseline given by the mean fixation impact introduced by Buscher in [36]. Their results showed that a relationship exists between EEG features and the users' attention to objects. In particular, the longer the subjects watched an object, the less the brain signal appeared disturbed. These results suggest that EEG features could be used to identify salient objects without considering the time users spend on them, as was done in previously mentioned methodologies.

3.4. Fusion of biometric resources or techniques

There is another group of studies in which sources of raw data are not fused directly, the fusion is performed in a more abstract way, such as resources or techniques.

The work of [41] examines the reliability of implicit feedback generated from clickthrough data in WWW search. Analyzing the users' decision process using eye tracking and comparing implicit feedback against manual relevance judgements, the authors conclude that clicks are informative but biased. While this makes the interpretation of clicks as absolute relevance judgements difficult, we show that relative preferences derived from clicks are reasonably accurate on average.

Furthermore, in [43], the authors studied the relationship between location typicality and efficiency in finding target Web objects for the case of online shops, online newspapers, and company web pages. For this, they performed an eye-tracking study in combination with data about mental models from their previous study [44]. They found that a typical object placement signified fewer fixations and a faster object finding. However, some Web objects were less sensitive to location typicality if they were more visually salient and conformed to user expectations in appearance.

In [37] Dimpfel and Morys used a combination of eye tracking and EEG to perform an objective assessment of five commercial websites. The eye-tracking device was used mainly for tracking gaze movements, while diverse quantitative features were obtained from the EEG. These features tried to measure users' attention and activation and then these reactions were compared with a typical survey. The results showed that the use of EEG features could be helpful in website analysis, but more studies are needed to confirm if this kind of research could be helpful in other scenarios, such as advertising.

In 2005, Li et al. performed an eye-tracking experiment to assess the users' web page viewing behavior, with the objective of finding some features to characterize it and proposing suggestions for web design. Their conclusions include the stability of the fixation duration distribution, the fixation duration dependency on some crowd factors and web page contents, and the fact that users' attentions were ordinal [45].

Djamasbi et al. examined the effect that pictures of faces had on the visual appeal, efficiency and trustworthiness of a web page, discovering that users believe that a page containing images of people's faces are more appealing and that it is easier to perform tasks in them, as opposed to those that do not contain them. Furthermore, the analysis revealed a strong positive correlation between trusting the informational content of a page and its visual appeal [46]. They use eye tracking to refine the results obtained in a preliminary study, which consisted of an online survey.

In [49] an EEG is used to determine a user's attention level and utilize it in three different ways. First, as a control mechanism, to control user interface elements such as menus or buttons. Second, to make the web browser responsive to the current attention level. Third, as a means for the web developer to control the user experience based on the level of attention paid by the user, thus creating attention-sensitive websites. They present implementation details for each of these and also explore issues in the system, together with the possibility of an EEG-based web standard.

The research of [70] used an eye tracker to track the eye-movement process of 42 college students when they were surfing websites with different manipulated levels of complexity and when completing simple and complex tasks respectively. The study examines how website complexity and task complexity jointly affect users' visual attention and behavior due to different cognitive loads. Results show that task complexity can moderate the effect of website complexity on users' visual attention and behavior. Specifically, when users conducted a simple task, fixation count and task completion time were at the highest level on the website with high complexity, while fixation duration was not significantly different on the websites with different complexity. However, when users conducted a complex task on a website with medium complexity, task completion time, fixation count, and fixation duration were all at their highest level. The findings provide guidelines for website managers and designers to maximize users' visual attention.

Based on an eye-tracking study of 30 subjects on 22 web pages from 11 popular web-sites, the study presented in [71] intends to explore the determinants of ocular behavior on a single web page whether it is determined: by individual differences of the subjects, different types of websites, the order of web pages being viewed, or the task at hand. The results indicate that the gender of subjects, the viewing order of a web page, and the interaction between page order and site type influences online ocular behavior. Task instruction did not significantly affect web viewing behavior. Scan-path analysis revealed that the complexity of web page design influences the degree of scan-path variation among different subjects on the same web page.

3.5. A conceptual framework for biometric data fusion in the web usage mining field

A practical model for information fusion was developed by the Joint Directors of Laboratories (JDL) Data Fusion Working Group of the United States Department of Defense. This model originally consisted of four levels that allow the identification of the processes, functions, categories of techniques and specific techniques applicable to data fusion [73,74].

We modified the original model adding a fifth level, following the same reasoning applied in different fields such as Intrusion Detection [90], Computer Security [75] and Opinion Mining [31]. Next the five levels are explained and related to each WUM process step.

- **Level 0 - Data refinement:** The lowest level of abstraction of data is viewed in this level, by filtering and calibrating it. In the WUM field, this level would be present in the data acquisition phase, considering web logs and eye-tracking recordings, for example. As in [31], this level is analogous to the first step of the architecture proposed by Dasarathy in 1997 [76], *Data In-Data Out Fusion*, where data is considered both an input and an output.
- **Level 1 - Object refinement:** This stage refers to the Preprocessing step of the WUM pipeline, which is the alignment of data into a common frame of reference or data structure, with the objective of using it all together at a subsequent point. According to Dasarathy's model, this is similar to the *Data In-Feature Out Fusion* concept. In this level, the raw data is preprocessed and transformed into features, for example EEG data could be decomposed into frequency bands and transformed into band power values to describe mental activity derived from diverse stimuli.
- **Level 2 - Situation refinement:** In this level, situational knowledge and awareness is obtained. That is, the aggregation of objects as sets of features, which have been put into a higher level of abstraction, permits a higher level of understanding of how the system is working. Based on Dasarathy's model, this is related to the *Feature In-Feature Out Fusion*, where features are supplied to methods, resulting in new sets of features to be used subsequently. In the context of this work, an example could be reducing the dimensionality of the problem, by performing diverse feature selection algorithms such as Principal Component Analysis (PCA).
- **Level 3 - Threat assessment:** The current situation is assessed based on juxtaposing the situational knowledge (of level 2) with previously determined templates. Under the model of [76], this is *Feature In-Decision Out Fusion*, due to the utilization of refined features into models or algorithms resulting in mixed decisions made by expert systems or humans. For instance, the content or structure of a website could be improved by knowing the emotional behavior patterns of the audience.
- **Level 4 - Resource management:** This last step considers a further refinement of the current situation. More thorough analysis could be made, in levels 0, 1 or 2, including acquiring new or additional data, for example.

Studying WUM implies the use of several sources of data, so it is actually an information fusion problem problem *per se*. Furthermore, using Biometric Information techniques for WUM analysis would mean the utilization of *more* data sources, so the primary importance of JDL model is that it could be used to describe the information fusion process in the particular subject we are exploring in this paper. Level 0 of the JDL model could be used to fuse diverse sources of data in the data acquisition step of the WUM problem, for instance the web log files and the web user eye movement in the screen. Likewise, level 1 could be used to transform the collected data into useful features as was done in the data

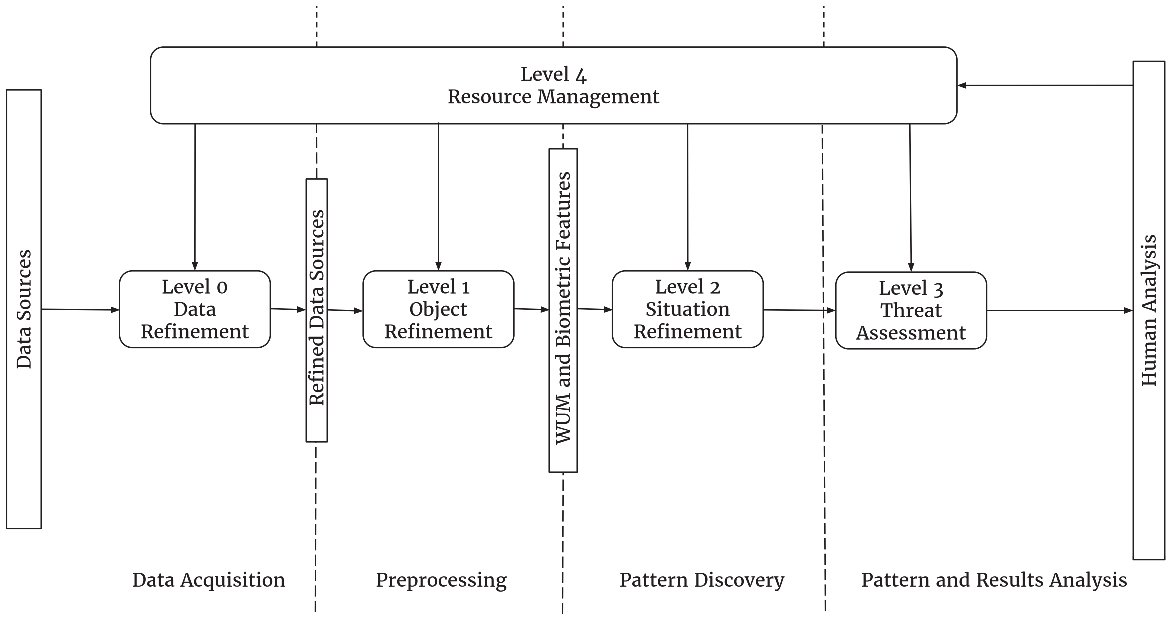


Fig. 7. Conceptual framework for biometric data fusion in WUM.

preprocessing phase, the collected data must to be transformed in feature vectors. Furthermore, pattern discovery matches with level 2, since refined features correspond to the outputs of the application of algorithms, whose results will be assessed in level 3, same as in the pattern and results analysis stage. All these WUM steps and the JDL model's levels could be revised at Level 4, for example, by including another source of data or changing the models for recognizing the patterns. Finally, the aggregation of the whole set would allow decisions to be made in order to utilize the knowledge gained regarding the Web user behavior studied (see Fig. 7).

4. Related work

In this section we present other surveys on the fields of Information fusion and on Web Biometrics. It is important to remark that while there are many surveys on Information fusion, there are much fewer surveys on Web Biometrics, which shows how new the field is.

4.1. Information fusion

Regarding Information fusion, the reader can be referred to [31], where the authors review the state of the art in Information fusion applied to opinion mining, and also review other 14 surveys in the field of Information fusion. Here we talk about other surveys not covered by that work. In [91] the authors, in the context of social big data, explore works on ontology-based information fusion, that is, integrating data with different conceptual structures and linguistic differences, and social network integration, that is, integrating mismatching information from more than one social network a user can have an account on. This paper discusses the benefits of fusing information when dealing with social data, characteristics of data from social media, approaches for combining information from different sources, applications of information fusion in social big data and a proposal of future directions on the field.

4.2. Web biometrics

To date, there are not many surveys on the use of biometrics for web usage mining. The work presented in [35] surveys the use of eye tracking in investigations of online searches. Authors analyze studies that reveal how users view the ranked results on a search engine results page, the relationship between the search result abstracts viewed and those clicked on, and whether gender, search task, or search engine influence these behaviors. They also discuss the use of eye tracking in studying online behaviors and its limited support for analyzing scan-paths, or sequences of eye fixations.

5. Conclusions

In this paper we have presented the biometric information fusion approach as a complement of traditional Web Usage Mining (WUM) techniques to understand the navigational behavior and preferences of Web users. We have briefly described the latter field, defined information fusion, reviewed some of the studies that have successfully implemented biometric information fusion in this specific context, and proposed a conceptual framework in order to guide the biometric information fusion process in WUM systems.

Since traditional WUM has always lacked real objectivity, the application of these new biometric and physiological perspectives is currently a positive step for such an important field. Creating systems which incorporate objective sources of knowledge such as eye trackers, electroencephalograms or more advanced brain imagery tools, will lead to better and deeper results. Fusing these with emotional or cognitive models will promote a finer understanding of Web user behavior and how to deliver what is really needed. Several previous studies have performed fusion without following any guidelines or frameworks, however the results are more realistic when they are incorporated.

Acknowledgments

The authors would like to acknowledge the continuing support of the Chilean Millennium Institute of Complex Engineering Systems (ICM: P-05-004-F, CONICYT: FBO16), and the Fondecyt Project 1160117.

References

[1] M. Spiliopoulou, Web usage mining for web site evaluation, Commun. ACM 43 (8) (2000) 127–134.
[2] P.E. Román, G. L'Huillier, J.D. Velásquez, Advanced Techniques in Web Intelligence - I, Springer Berlin Heidelberg, Berlin, Heidelberg, 2010, pp. 143–165.

[3] B. Mobasher, H. Dai, T. Luo, M. Nakagawa, Effective personalization based on association rule discovery from web usage data, in: Proceedings of the 3rd International Workshop on Web Information and Data Management, ACM, 2001, pp. 9–15.
[4] M. Aldekhail, Application and significance of web usage mining in the 21st century: a literature review, Int. J. Comput. Theory Eng. 8 (1) (2016) 41.
[5] J.V. Haxby, M.I. Gobbini, M.L. Furey, A. Ishai, J.L. Schouten, P. Pietrini, Distributed and overlapping representations of faces and objects in ventral temporal cortex, Science 293 (5539) (2001) 2425–2430.
[6] B. Mobasher, Web Data Mining: Exploring Hyperlinks, Contents, and Usage Data, Springer Berlin Heidelberg, Berlin, Heidelberg, 2007, pp. 449–483.
[7] P.H. Foo, G.W. Ng, High-level information fusion: an overview, J. Adv. Inf. Fusion 8 (1) (2013) 33–72.
[8] H. Boström, S.F. Andler, M. Brohede, R. Johansson, A. Karlsson, J. Van Laere, L. Niklasson, M. Nilsson, A. Persson, T. Ziemke, On the definition of information fusion as a field of research, Tech. Rep. HS-IKI-TR-07-006 (2007).
[9] F.M. Facca, P.L. Lanzi, Mining interesting knowledge from weblogs: a survey, Data Knowl. Eng. 53 (3) (2005) 225–241.
[10] R. Kosala, H. Blockeel, Web mining research: a survey, SIGKDD Explor. Newsl. 2 (1) (2000) 1–15.
[11] F. Johnson, S.K. Gupta, Article: web content mining techniques: a survey, Int. J. Comput. Appl. 47 (11) (2012) 44–50.
[12] P.R. Kumar, A.K. Singh, Web structure mining: exploring hyperlinks and algorithms for information retrieval, Am. J. Appl. Sci. 7 (6) (2010) 840–845.
[13] J. Srivastava, R. Cooley, M. Deshpande, P.-N. Tan, Web usage mining: discovery and applications of usage patterns from web data, SIGKDD Explor. Newsl. 1 (2) (2000) 12–23.
[14] D. Pierrakos, G. Paliouras, C. Papatheodorou, C.D. Spyropoulos, Web usage mining as a tool for personalization: a survey, User Model. User Adapt. Interact. 13 (4) (2003) 311–372.
[15] A. Nanopoulos, Y. Manolopoulos, M. Zakrzewicz, T. Morzy, Indexing web access-logs for pattern queries, in: Proceedings of the 4th International Workshop on Web Information and Data Management, ACM, 2002, pp. 63–68.
[16] K. Suneetha, R. Krishnamoorthi, Identifying user behavior by analyzing web server access log file, IJCSNS Int. J. Comput. Sci. Netw. Security 9 (4) (2009) 327–332.
[17] T. Hussain, S. Asghar, N. Masood, Web usage mining: a survey on preprocessing of web log file, in: Information and Emerging Technologies (ICIET), 2010 International Conference on, 2010, pp. 1–6.
[18] M. Munk, J. Kapusta, P. Švec, Data preprocessing evaluation for web log mining: reconstruction of activities of a web visitor, Procedia Comput. Sci. 1 (1) (2010) 2273–2280.
[19] R. Cooley, B. Mobasher, J. Srivastava, Data preparation for mining world wide web browsing patterns, Knowl. Inf. Syst. 1 (1) (1999) 5–32.
[20] D. Tanasa, B. Trousse, Advanced data preprocessing for intersites web usage mining, IEEE Intell. Syst. 19 (2) (2004) 59–65.
[21] J. Yao, V.V. Raghavan, Z. Wu, Web information fusion: a review of the state of the art, Inf. Fusion 9 (4) (2008) 446–449.
[22] K. Chen, Large-scale deep web integration: exploring and querying structured data on the deep web, in: Tutorial at In Proceedings of Proceedings of the ACM SIGMOD International Conference on Management of Data, 2006, pp. 27–29.
[23] A. Halevy, A. Rajaraman, J. Ordille, Data integration: the teenage years, in: Proceedings of the 32Nd International Conference on Very Large Data Bases, in: VLDB '06, VLDB Endowment, 2006, pp. 9–16.
[24] R. Kambhampati, C. Knoblock, Information integration on the web, Tut. AAAI 07 (2007).
[25] M. Perkowitz, O. Etzioni, Towards adaptive web sites: conceptual framework and case study, Comput. Netw. 31 (11) (1999) 1245–1258.
[26] J.D. Velásquez, P.A. Estévez, H. Yasuda, T. Aoki, E. Vera, Intelligent web site: understanding the visitor behavior, in: Knowledge-Based Intelligent Information and Engineering Systems, Springer, 2004, pp. 140–147.
[27] J.D. Velásquez, V. Palade, Adaptive web sites: a knowledge extraction from web data approach, vol. 170, Ios Press, 2008.
[28] B. Mobasher, R. Cooley, J. Srivastava, Automatic personalization based on web usage mining, Commun. ACM 43 (8) (2000) 142–151.
[29] M. Eirinaki, M. Vazirgiannis, Web mining for web personalization, ACM Trans. Internet Technol. 3 (1) (2003) 1–27.
[30] S.S. Anand, B. Mobasher, Intelligent techniques for web personalization, in: Proceedings of the 2003 International Conference on Intelligent Techniques for Web Personalization, in: ITWP'03, Springer-Verlag, Berlin, Heidelberg, 2005, pp. 1–36.
[31] J.A. Balazs, J.D. Velásquez, Opinion mining and information fusion: a survey, Inf. Fusion 27 (2016) 95–110.
[32] D.L. Hall, J. Llinas, An introduction to multisensor data fusion, Proc. IEEE 85 (1) (1997) 6–23.
[33] L. Wald, A European proposal for terms of reference in data fusion, in: Commission VII Symposium "Resource and Environmental Monitoring", 1998, pp. 651–654. Budapest, Hungary
[34] B. Khaleghi, A. Khamis, F.O. Karray, S.N. Razavi, Multisensor data fusion: a review of the state-of-the-art, Inf. Fusion 14 (1) (2013) 28–44.
[35] L. Lorigo, M. Haridasan, H. Brynjarsdóttir, L. Xia, T. Joachims, G. Gay, L. Granka, F. Pellacini, B. Pan, Eye tracking and online search: lessons learned and challenges ahead, J. Am. Soc. Inf. Sci. Technol. 59 (7) (2008) 1041–1052.
[36] G. Buscher, E. Cutrell, M.R. Morris, What do you see when you're surfing?: using eye tracking to predict salient regions of web pages, in: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, in: CHI '09, ACM, New York, NY, USA, 2009, pp. 21–30.
[37] W. Dimpfel, A. Morys, Quantitative objective assessment of websites by neurocode-tracking in combination with eye-tracking, J. Behav. Brain Sci. 4 (8) (2014).
[38] P. Loyola, J.D. Velásquez, Characterizing web user visual gaze patterns: a graph theory inspired approach, in: International Conference on Brain Informatics and Health, Springer, 2014, pp. 586–594.
[39] H. Lee, S. Seo, A comparison and analysis of usability methods for web evaluation: The relationship between typical usability test and bio-signals characteristics (eeg, ecg), in: Proc. of 2010 DRS (Design Research Society) Montreal Conference, 2010, pp. 893–904.
[40] V. do Amaral, L.A. Ferreira, P.T. Aquino, M.C.F. de Castro, Eeg signal classification in usability experiments, in: Biosignals and Biorobotics Conference (BRC), 2013 ISSNIP, 2013, pp. 1–5.
[41] T. Joachims, L. Granka, B. Pan, H. Hembrooke, G. Gay, Accurately interpreting clickthrough data as implicit feedback, in: Proceedings of the 28th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, in: SIGIR '05, ACM, New York, NY, USA, 2005, pp. 154–161.
[42] G. Buscher, S.T. Dumais, E. Cutrell, The good, the bad, and the random: an eye--tracking study of ad quality in web search, in: Proceedings of the 33rd International ACM SIGIR Conference on Research and Development in Information Retrieval, in: SIGIR '10, ACM, New York, NY, USA, 2010, pp. 42–49.
[43] S.P. Roth, A.N. Tuch, E.D. Mekler, J.A. Bargas-Avila, K. Opwis, Location matters, especially for non-salient features: an eye-tracking study on the effects of web object placement on different types of websites, Int. J. Hum. Comput. Stud. 71 (3) (2013) 228–235.
[44] S.P. Roth, P. Schmutz, S.L. Pauwels, J.A. Bargas-Avila, K. Opwis, Mental models for web objects: where do users expect to find the most frequent objects in online shops, news portals, and company web pages? Interact. Comput. 22 (2) (2010) 140–152.
[45] Q. Li, L. Sun, J. Duan, Web page viewing behavior of users: an eye-tracking study, in: Proceedings of ICSSSM '05. 2005 International Conference on Services Systems and Services Management, 2005., vol. 1, 2005, pp. 244–249.
[46] S. Djamasbi, M. Siegel, T. Tullis, Efficiency, trust, and visual appeal: usability testing through eye tracking, in: Proceedings of the Forty-Third Annual Hawaii International Conference on System Sciences (HICCS), Computer Society Press, 2010, pp. 1–10.
[47] J. Jadue, G. Slanzi, L. Salas, J.D. Velásquez, Web user click intention prediction by using pupil dilation analysis, in: 2015 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT), 2015, pp. 433–436.
[48] D. Bansal, J. Bose, A. Kumar, Eeg based detection of area of interest in a web page, in: Advances in Computing, Communications and Informatics (ICACCI), 2015 International Conference on, 2015, pp. 1320–1325.
[49] J. Bose, A. Singhai, A.A. Patankar, A. Kumar, Attention sensitive web browsing, CoRR abs/1601.01092 (2016).
[50] J.H. Goldberg, A.M. Wichansky, Eye tracking in usability evaluation: a practitioner's guide, Elsevier Sci. (2003) 493–516.
[51] Tobii pro glasses 2 wearable eye tracker, URL http://www.tobiipro.com/product-listing/tobii-pro-glasses-2. Accessed: February 10, 2017.
[52] Biosemi eeg ecg emg bspm neuro amplifier electrodes, URL http://www.biosemi.com/products.htm. Accessed: February 10, 2017.
[53] Emotiv epoc - 14 channel wireless eeg headset, Accessed: February 10, 2017. URL https://www.emotiv.com/epoc.
[54] A. Poole, L.J. Ball, Eye tracking in human-computer interaction and usability research: currents status and future prospects, Encycl. Human Comput. Interact. 1 (2006) 211–219.
[55] G. Slanzi, J.A. Balazs, J.D. Velásquez, Combining eye tracking, pupil dilation and eeg analysis for predicting web users click intention, Inf. Fusion 35 (2017) 51–57.
[56] E. Cutrell, Z. Guan, What are you looking for?: an eye-tracking study of information usage in web search, in: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, ACM, 2007, pp. 407–416.
[57] M. Teplan, Fundamentals of eeg measurement, Measure. Sci. Rev. 2 (2) (2002) 1–11.
[58] M. Teplan, Fundamentals of eeg measurement, Measure. Sci. Rev. 2 (2) (2002) 1–11.
[59] N.-Y. Liang, P. Saratchandran, G.-B. Huang, N. Sundararajan, Classification of mental tasks from eeg signals using extreme learning machine, Int. J. Neural Syst. 16 (01) (2006) 29–38.
[60] N. Zulkifli, M. Othman, N. Kamal, A review: fundamental and applications of functional magnetic resonance imaging (fmri) on brain learning activities, Adv. Appl. Inf. Sci. (2012) 85–90.
[61] S. Ghebreab, A. Smeulders, P. Adriaans, Predicting brain states from fmri data: incremental functional principal component regression, in: J.C. Platt, D. Koller, Y. Singer, S.T. Roweis (Eds.), Advances in Neural Information Processing Systems 20, Curran Associates, Inc., 2008, pp. 537–544.
[62] J. Gwizdka, Y. Zhang, Towards inferring web page relevance an eye-tracking study, iConference 2015 Proc. (2015).
[63] K. Wise, H.J. Kim, J. Kim, The effect of searching versus surfing on cognitive and emotional responses to online news, J. Media Psychol. 21 (2) (2009) 49–59.
[64] J.D. Velásquez, Web site keywords: a methodology for improving gradually the web site text content, Intell. Data Anal. 16 (2) (2012) 327–348.
[65] L.E. Dujovne, J.D. Velásquez, Design and implementation of a methodology for identifying website keyobjects, in: Proceedings of the 13th International Conference on Knowledge-Based and Intelligent Information and Engineering Systems, KES, 2009, pp. 301–308.

[66] S.A. Ríos, J.D. Velásquez, H. Yasuda, T. Aoki, Web site improvements based on representative pages identification, in: Australasian Joint Conference on Artificial Intelligence, Springer Berlin Heidelberg, 2005, pp. 1162–1166.
[67] J.D. Velásquez, L.E. Dujovne, G. L'Huillier, Extracting significant website key objects: a semantic web mining approach, Eng. Appl. Artif. Intell. 24 (8) (2011) 1532–1541. Semantic-based Information and Engineering Systems.
[68] J.D. Velásquez, Combining eye-tracking technologies with web usage mining for identifying website keyobjects, Eng. Appl. Artif. Intell. 26 (56) (2013) 1469–1478.
[69] P. Loyola, G. Martinez, K. Muñoz, J.D. Velásquez, P. Maldonado, A. Couve, Combining eye tracking and pupillary dilation analysis to identify website key objects, Neurocomputing 168 (2015) 179–189.
[70] Q. Wang, S. Yang, M. Liu, Z. Cao, Q. Ma, An eye-tracking study of website complexity from cognitive load perspective, Decis. Support Syst. 62 (2014) 1–10.
[71] B. Pan, H.A. Hembrooke, G.K. Gay, L.A. Granka, M.K. Feusner, J.K. Newman, The determinants of web page viewing behavior: an eye-tracking study, in: Proceedings of the 2004 Symposium on Eye Tracking Research & Applications, in: ETRA '04, ACM, New York, NY, USA, 2004, pp. 147–154.
[72] G. Slanzi, C. Aracena, J.D. Velásquez, Eye tracking and eeg features for salient web object identification, in: International Conference on Brain Informatics and Health, Springer, 2015, pp. 3–12.
[73] F.E. White, Data Fusion Lexicon, Technical Report, DTIC Document, 1991. Accessed: February 10, 2017.
[74] D.L. Hall, J. Llinas, An introduction to multisensor data fusion, Proc. IEEE 85 (1) (1997) 6–23.
[75] I. Corona, G. Giacinto, C. Mazzariello, F. Roli, C. Sansone, Information fusion for computer security: state of the art and open issues, Inf. Fusion 10 (4) (2009) 274–284.
[76] B.V. Dasarathy, Sensor fusion potential exploitation-innovative architectures and illustrative applications, Proc. IEEE 85 (1) (1997) 24–38.
[77] F. Ricci, L. Rokach, B. Shapira, Introduction to Recommender Systems Handbook, Springer US, Boston, MA, 2011, pp. 1–35.
[78] R. Forsati, M.R. Meybodi, A. Rahbar, An efficient algorithm for web recommendation systems, in: 2009 IEEE/ACS International Conference on Computer Systems and Applications, 2009, pp. 579–586.
[79] D. Parra, S. Sahebi, Recommender systems: sources of knowledge and evaluation metrics, in: Advanced Techniques in Web Intelligence-2, Springer, 2013, pp. 149–175.
[80] X. Jin, Y. Zhou, B. Mobasher, A maximum entropy web recommendation system: combining collaborative and content features, in: Proceedings of the Eleventh ACM SIGKDD International Conference on Knowledge Discovery in Data Mining, in: KDD '05, ACM, New York, NY, USA, 2005, pp. 612–617.
[81] K. Uludag, D.J. Dubowitz, R.B. Buxton, Basic principles of functional mri, in: R. Edelman, J. Hesselink, M. Zlatkin, J. Crues (Eds.), Clinical magnetic resonance imaging. 3rd edition, Elsevier, 2006, pp. 249–287.
[82] D. Shanmugapriya, G. Padmavathi, A survey of biometric keystroke dynamics: approaches, security and challenges, arXiv preprint arXiv:0910.0817(2009).
[83] F. Bergadano, D. Gunetti, C. Picardi, User authentication through keystroke dynamics, ACM Trans. Inf. Syst. Security (TISSEC) 5 (4) (2002) 367–397.
[84] H. Jasper, Report of the committee on methods of clinical examination in electroencephalography, Electroencephalogr. Clin. Neurophysiol. 10 (1958) 370–375.
[85] M. Rybnik, P. Panasiuk, K. Saeed, User authentication with keystroke dynamics using fixed text, in: 2009 International Conference on Biometrics and Kansei Engineering, 2009, pp. 70–75.
[86] D.S. Wooding, Fixation maps: quantifying eye-movement traces, in: Proceedings of the 2002 Symposium on Eye Tracking Research & Applications, ACM, 2002, pp. 31–36.
[87] S.T. de Magalhaes, K. Revett, H.M.D. Santos, Password secured sites - stepping forward with keystroke dynamics, in: International Conference on Next Generation Web Services Practices (NWeSP'05), 2005, pp. 1–6.
[88] I. Traore, I. Woungang, M.S. Obaidat, Y. Nakkabi, I. Lai, Combining mouse and keystroke dynamics biometrics for risk-based authentication in web environments, in: 2012 Fourth International Conference on Digital Home, 2012, pp. 138–145.
[89] B. Kemp, A. Värri, A.C. Rosa, K.D. Nielsen, J. Gade, A simple format for exchange of digitized polygraphic recordings, Electroencephalogr. Clin. Neurophysiol. 82 (5) (1992) 391–393.
[90] T. Bass, Intrusion detection systems and multisensor data fusion, Commun. ACM 43 (4) (2000) 99–105.
[91] G. Bello-Orgaz, J.J. Jung, D. Camacho, Social big data: recent achievements and new challenges, Inf. Fusion 28 (2016) 45–59.