# BIOMETRICS AS AN INTELLIGENT PART OF THE B2C ENVIRONMENT

Milomir Tatović,
Saša Adamović,
Milan Milosavljević

Singidunum University,
32 Danijelova Street, Belgrade, Serbia

Abstract:
This paper discusses the possibilities of behavioural biometrics in content analysis of the web site and the end user. By collecting data through the human - computer interface (mouse, keyboard or touch screen), it is possible to obtain sufficient amount of information to reliably identify users. With precise identification it would be possible to develop an adaptive web site model, which would represent a great contribution to the B2C environment.

Key words:
biometrics, mouse tracking, B2C.

## 1. INTRODUCTION

Look in the eyes, crossed hands or hands in the pockets during the conversation reveal much more than words. Candies in markets placed at the height of children eye sight and traditional painting of fast food restaurants with colours that cause hunger are some of the marketing tips that traders do in order to trick the customers. These tricks and reading body language are our real everyday life. However, the question is what happens in virtual life, where you cannot see or feel, and sometimes not even hear the reaction of the other side, where it is impossible to know if this person even exists or is it just a virtual world person - a bot.

Obviously, the question arises whether we can use real life psychological methods which are being used very successfully, and apply them in virtual reality. First question that needs to be answered is: Can we say with certainty that we are dealing with human? The answer to this question is quite interesting. Many programmers are trying to come up with an efficient system which will be able to distinguish human from each bot. On the other side, there is a large number of programmers who are trying to take the full advantage of artificial intelligence and create a machine-bot that is difficult to recognize.

Such systems, which allow the detection between a human or a robot are based on the elements of nonverbal communication. A computer program cannot detect waving hands but it can detect things like uncontrolled mouse movement, a large number of simultaneous clicks or many syntax errors. Owning to this we can claim with certainty that here is a human behind the computer, and we can also easily evaluate mental condition or characteristics of the virtual reality user.

Correspondence:
Milomir Tatović

e-mail:
milomir@tatovic.com



DOI: 10.15308/Sinteza-2016-161-164

## 2. MONITORING SYSTEMS

As the greatest representative of virtual reality, web requires more end-users analysis. In real life, shop owners know exactly how to arrange their articles and that's how the Web developers want to make an appropriate Web page in virtual world. Here, the user analysis is more complex. Primarily, it is necessary to attract the attention of a potential buyer and to distribute items on a Web page exactly in the way the customer would prefer. In order to obtain such information from customers, it is necessary to pay attention to their activities on the Web, and on the page itself. In addition, time of expensive solutions is over. Web is full of free or "almost" free content. A large number of ads on websites are becoming an important source of financing solutions. Therefore, it is necessary to place advertising exactly where it will certainly be noticed by the end user.

The most common system of analysis, as well as in real life and on the Web, is a test system A/B[1]. In this way, we get reply from potential users or end users prototypes. Yet, it is sometimes impossible to provide the conditions necessary for testing. Then and it is necessary to monitor the activities of real users, and adjust the web page to their needs or to grow seamlessly up to the desired activity. This kind of a system requires a great source of data. In human-computer interaction, mouse movement is an ideal source. [2] Following the path of the cursor or places where the cursor is generally retained, the way how you scroll or intensity of the necessary and random clicks, can give us a lot of information about the site visitor. [3] Notable difference between visitors will be discussed in the next section of this document.

The number of available systems is not great. The problem is certainly a large amount of data generated every moment. However, with the constant growth of processor power and high availability of high-speed memory devices, these problems become less important. Therefore, it is nowadays possible to monitor eye view besides the mouse movement. When we combine mouse movement and eye monitoring of the end user, we get better data to analyse user interactions. In a study conducted by Kerry Rodden and Xin Fu [4], a correlation between the cursor position and views has been demonstrated. Rodden also presented four different approaches of using the mouse. The first is neglecting the mouse during reading the text, the second and third indicate that users use the mouse by tracking the mouse horizontally or vertically to read the text and the fourth shows that some users mark interesting results when reading the text. Guo and Agichtein have proven in their study that they can predict with 77% accuracy where the user was looking by tracking the cursor.[5]

If we decide to follow a viewpoint on a website, we need to keep in mind that it is necessary that the user has a camera and has agreed to be filmed during the site visit. Namely, if we follow only the moves of the mouse, clicks, typing, or scroll, user does not know that he is being monitored and there is no need for performing an additional action. This way, we can provide high-quality data for analysis. With the results obtained from these studies, we can claim that we don't need monitoring views for this kind of analysis.

## 3. DATA COLLECTION AND RECONSTRACTION

Data collection systems can be desktop programs or Web-based applications. Desktop programs must be installed by the user, so they are not eligible for the analysis. A web-based system was made for the purpose of this study. Interactive visitor data interactions are collected directly from the browser with the help of java script functions and sent and stored in the database. In this way, the end user does not know that his interaction was monitored nor the functionality of the site was disrupted. This way monitoring result obtained can be considered relevant. Besides the exact URL, mouse movements, clicks, scrolls and typing on the keyboard, other user information is also obtained. Those are IP address, country and city, the name and version of the operating system and browser, also unique cookie that is created when you first visit the site, which is later stored in the memory of the user. Creating a cookie allows us to know with certainty that the same user has returned. When creating this system, in particular, we made sure that we enable a realistic reconstruction of the recorded content. For this reason, one of the JavaScript functions is responsible for the source site collection and works on the principle of mapping elements, and collected clicks or mouse movements are not based on pixel values, but on the percentage of the actual position of some HTML DOM elements. [6] Using the hash function on the mapped DOM, we can easily monitor any changes that occur on the site. Apart from that we have provided simple analysis of website content because when you add a new DOM element and change the Web site appearance, we still retain all the previous information about the previous items. With the collection of data on the size of the entire page (the page size), the size of the visible part of the site (the window size), and the constant tracking

of the scroll values, we are able to know the exact part of the site which was viewed by the visitors.

The data collected needs to be reproduced. For the purpose of this paper, several different types of reconstruction were realized. Reconstruction is divided into three parts, in which data were taken into account.

*Reconstruction based on visits*

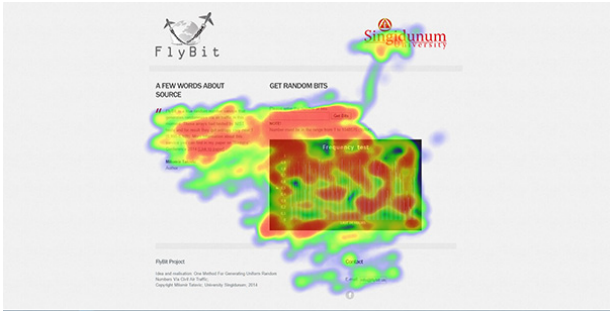


Figure 1. Heatmap of Mouse move


Figure 2. Heatmap of clicks

If we want to analyse the visit of a website, we need the reconstruction based on all the visits of that URL for the selected period. This type of analysis is best seen through a Heat map. The first heat map is created based on the scroll values and resolutions users had while navigating the site. It shows the visibility of the site, or precisely the number of users that has seen a certain part of the site. Second heat map shows a map of clicks on a web page. As already stated, all data is collected on the basis of mapped DOM, as in the case of Heat Map e-clicks makes us easy to show, because we don't need to take care of the different visitor's resolutions, and therefore the position of elements on the site, but only about those elements that were clicked on. Finally, the third heat map, the map that represents mouse movement. This map and map of clicks work on a similar principle, only mouse movement map takes a larger amount of data.

*Reconstruction based on one visit*


Figure 3. Single mouse tracking

This type of reproduction allows us to follow a specific user in real time or after completing a visit to the site. With the help of data collected, it is possible to display all visible changes that the visitor had, and even those made by some external script. This form of reconstruction was necessary primarily because of the following type.

*Reconstruction based on the dependence between one and all visits*

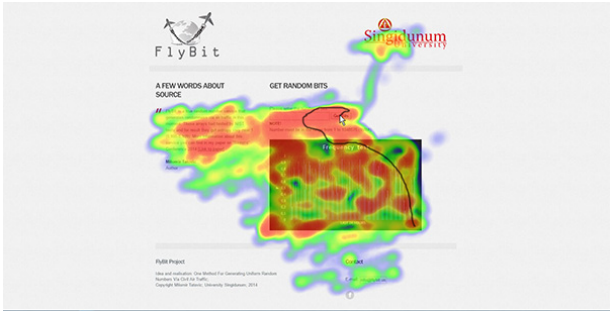


Figure 4. Single mouse tracking on mouse move heatmap

This type of reconstruction was necessary for continuing the paper. Here we can see how a visitor uses the site (real time or later after the visit), while in the background we can show a Heat map of all visits of all users or all visits of one particular user. In addition to this, it is possible to analyse the interaction with certain elements on the site. For example, we can see how often a certain button is clicked on the site. Some users always clicks on it when they enter the site, but on average basis, the button is clicked by every third visitor. These data are of great importance for further work.

## 4. TRACKING WITH BIOMETRIC ELEMENTS

Systems such as described above are based on monitoring of a particular user only via cookie. Such an approach is good if users do not delete cookies on their own or if they disable deleting cookies in browser settings. The question is why would someone do that? The largest number of ads that appear on websites are specifically based on cookies. As the user surfs the internet, the browser assigns a specific cookie on advertisement web pages. Later when the user visits a site where there are advertisements he gets the advertisements about a specific product he searched for. If the user does not have a cookie enabled like in the most cases, the advertisement cannot be displayed because there is no saved cookie data.

If we take such a system and add biometric knowledge, we get a complex system. It is known that it's very difficult to imitate anyone, so it is ideal for biometric systems such as this one. In those kind of systems, it is necessary to conclude whether the user is human or bot and later analyse if that user has already visited the web page.

The data we collect are essentially the same or substantially similar. Operating system or browser version are not so important. IP address, can mean a lot, but still cannot say anything. However, the mouse movement and clicks, like we discussed earlier give us a lot of information about visitors. Let's define the initial moment as the time when the user moves the mouse across the web site. In the first testing, we can observe the trajectory of mouse movement. If we decide to keep track of user in real time, or track a recorded session, we can activate the heatmap of mouse movement and thus immediately see if this session stands out from the rest. If this is the user who has visited previously, we can tell the probability based on the collected data whether it is the same user or not. The trajectory of the mouse movement and the cursor input at the visible part of the page carries a large amount of information. After the comparative tests it can easily be shown that the angle of the trajectory path is the best option for detecting the similar mouse movements. Straight or diagonal movement are most often characteristics of bot scripts. As shown in Figure 5., if we follow the new entry of the old visitors, we can detect similarities in movement. In addition to mouse movements, the second biometric characteristic is certainly the way a person is clicking or scrolling.

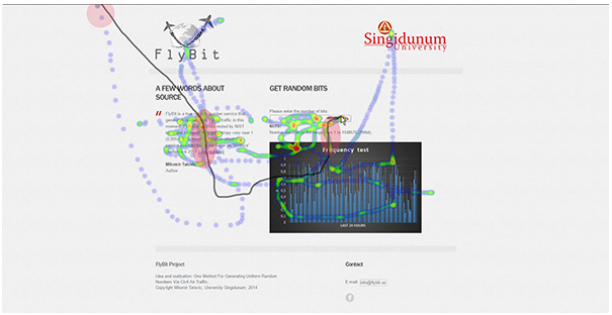


Figure 5. Live Mouse tracking on same Visitor mouse move heatmap

## 5. CONCLUSION

The system can provide more effective site analytics with the help of biometrics. In case of e-commerce, which is growing popularity, it can facilitate business to dealers and to customers and also secure sites from malicious impact. The development of this system as a proposal for future work should be undertaken more detailed visitor analysis, with the use of the psychological elements of real life, which would also facilitate research and save time.

## REFERENCES

[1] Optimizely – A/B testing. Retrieved March 31, 2016, from https://www.optimizely.com/case-studies/.

[2] Navalpakkam, V., Churchill, E. (2012). Mouse tracking: measuring and predicting users' experience of web-based content. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '12). ACM, New York, NY, USA, 2963-2972.

[3] Lee, G. Zhongming Chen, (2006). Investigating the Differences in Web BrowsingBehaviour of Chinese and European User Using Mouse Tracking. N. Aykin (Ed.):Usability and Internationalization, Part I, HCII 2007, LNCS 4559, pp. 502–512, 2007. Springer-Verlag Berlin Heidelberg.

[4] K. Rodden, X. Fu, A. Aula, and I. Spiro. Eye-mouse coordination patterns on web search results pages. Ext. Abstracts CHI '08, 2997–3002.

[5] Qi Guo, E. Agichtein. Towards Predicting Web Searcher Gaze Position from Mouse Movements. CHI 2010, April 10–15, 2010, Atlanta, Georgia, USA. ACM 978-1-60558-930-5/10/04.

[6] Document Object Model (DOM). Retrieved March 31, 2016, from https://www.w3.org/DOM/