Eye Can Tell: On the Correlation between Eye Movement and Phishing Identification

Daisuke Miyamoto, Gregory Blanc, and Youki Kadobayashi

[1] Information Technology Center
The University of Tokyo
2-11-16 Yayoi, Bunkyo-ku, Tokyo, 113-8658 JAPAN
daisu-mi@nc.u-tokyo.ac.jp
[2] Institut Mines-Télécom/Télécom SudParis
CNRS UMR 5157 SAMOVAR
9 rue Charles Fourier, 91011 Évry, FRANCE
gregory.blanc@telecom-sudparis.eu
[3] Graduate School of Information Science
Nara Institute of Science and Technology
8916-5 Takayama, Ikoma, Nara, 630-0192 JAPAN
youki-k@is.aist-nara.ac.jp

Abstract. It is often said that the eyes are the windows to the soul. If that is true, then it may also be inferred that looking at web users' eye movements could potentially reflect what they are actually thinking when they view websites. In this paper, we conduct a set of experiments to analyze whether user intention in relation to assessing the credibility of a website can be extracted from eye movements. In our within-subject experiments, the participants determined whether twenty websites seemed to be phishing websites or not. We captured their eye movements and tried to extract intention from the number and duration of eye fixations. Our results demonstrated the possibility to estimate a web user's intention when making a trust decision, solely based on the user's eye movement analysis.

Keywords: Phishing, Cognitive Psychology, Eye-Tracking

1 Introduction

Phishing is a fraudulent activity defined as the acquisition of personal information by tricking an individual into believing the attacker is a trustworthy entity [1]. Phishing attackers usually lure people through the use of a phishing email, which appears to be sent by a legitimate corporation. The attackers then attract the email recipients to a phishing site, which is a replica of an existing web page, to fool them into submitting personal, financial, and/or password data.

In this paper, we propose a method to estimate which user is going to be victim of phishing on the basis of their eye movement analysis. According to the

theory of mind [8], eye movements are different between users with a motivation to find any particular objects and users without In the context of phishing identification, expert users may gain information from intentionally looking at the browser's address bar, and evaluating this information based on their knowledge. By contrast, novice users would look at the address bar with no particular motivation, simply because they are unable to evaluate this piece of information due to their lack of knowledge. Instead, novice users may intend to gain information from the website's contents even if the contents may not give any meaningful indications with regards to phishing identification.

One major obstacle to that proposal is that the eye movements are affected by many factors - age, eyesight, stress, knowledge, level of vigilance, awareness, and familiarity with the website as well as user's intention. We conjecture that eye movement certainly informs on the user's intention. Our challenge is to assess the user's intention based on the extracted information even if other factors are involved. The results may lead to estimating whether a user is likely to fall victim to phishing.

This paper therefore assesses this hypothesis with a participant based experiment in which 23 participants have their eye movements monitored while taking a test where they need to determine which websites are phish sites among twenty samples and provide their decision's criteria. Based on our experiment, it might be reasonable to consider that the analysis of eye movement is feasible for estimating users' both intention and decision.

The rest of the paper is organized as follows. Section 2 provides theoretical background on extracting human implicit intention, and Section 3 proposes our method which aims at estimating user intention while assessing the credibility of the websites. Section 4 explains the conditions for our experiments, and Section 5 evaluates the performance of our proposed method. Section 6 shows our follow-up study, and finally, Section 7 summarizes our contributions.

2 Related Work

Cognitive psychology is the study of the relationship between internal mental processes and observable behavior. In this paper, observation is carried out with respects to the criteria formulated by Groojten [2] for the evaluation of cognitive methods for supporting operators. These criteria are as follows:

- **Sensitivity to workload changes.** We need to employ the behavioral observation methods that can estimate the internal mental model. The methods might also leverage the collected information regardless of the Fear of Negative Evaluation (FNE) [11]; observations are often affected by FNE, in which some people will attempt to conceal their errors. In fact, disclosing mistakes often damage their own self-image and professional standing.
- **Obtrusiveness for the operator.** The observation should not take much effort to start collecting data or disturb the handling of people during the tasks performance. Furthermore, people will not carry implants, needles or other devices which may hurt them in any way.

– **Availability of equipment.** The observation should employ the method which is easily applicable to people. Within the context of phishing prevention, the method should be available while users are browsing. Non-contact devices might be preferred.

In this paper, we decided to employ eye movement-based observations since it meets the above requirements. Brain activity, heart measure, and blood pressure are feasible due to the sensitivity to workload changes, but they tend to require much more obtrusive monitoring devices. By contrast, Facial expression [5] and Gesture recognition [3] were often affected by FNE.

According to Leigh et al., eye movement is generally classified into four categories, namely Saccades, Smooth pursuit movements, Fixations, and Vestibulo-ocular reflexes [6]. Saccades are rapid, ballistic movements of the eyes that abruptly change the point of fixation. Tokuda showed that mental workload, the indicator of how mentally busy a person is, can be estimated from Saccadic intrusions [10]. By contrast, Smooth pursuit movements are slow and continuous eye movements that are used to track an object in motion with central vision in order to maintain a clear and continuous perception of it, and they are deficient in schizophrenia patients [9]. Fixation is the eye movement that maintains the visual gaze on a single location, and vestibulo-ocular reflexes stabilize gaze during movement.

Regarding eye fixation, prior studies [4, 7] contributed to show that there may be correlation between eye fixation and intention. The intention refers to an idea or plan of what a person is going to do. The theory of mind states that a person has a natural way to predict, represent and interpret intention expressed explicitly or implicitly [8]. A person expresses explicit intentions using different sequences of actions. For example, during an interaction, a person tends to express intention explicitly through speech, gesture, and facial expression. By contrast, implicit human intentions are subtle, vague and otherwise often difficult to interpret. Since the explicit expression alone may not be enough to understand the intention of a person, it is critical to understand the implicit intention.

According to [4, 7], the implicit intentions can be identified through the following biomedical signals during a visual stimulus.

– *Navigational intention* refers to an idea or plan of a person to find any object in a visual input without a particular motivation.
– *Informational intention* refers to an idea or plan of a person to find a particular object of interest or to behave with a motivation.

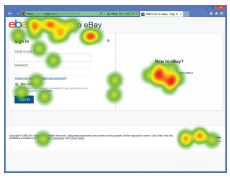
The authors have built classifiers for identifying intentions based on Support Vector Machine (SVM), and observed that there were positive correlations between the intentions and eye movement patterns.

3 Proposal

In this paper, we evaluate the feasibility of estimating the user's decision to trust or not the websites, when assessing the credibility of the websites, based

(a) Novice users (b) Experts

Fig. 1: Eye-tracking in a phishing site



(a) Novice users (b) Experts

Fig. 2: Eye-tracking in a legitimate site

on the user's eye movement patterns. The key idea is to apply the analysis of eye fixation, an established technique in the research domain of cognitive psychology, in order to improve security.

We consider that eyes are suitable to monitor personal mental processes. As we mentioned in Section 2, eyes can give information on whether a person has the intention to look for something. In the context of phishing prevention, a web user is presented with security indicators and web contents displayed in the browser. Our assumption is that experts have the intention to check security indicators rather than web contents, while novices would do the opposite.

Our primary motivation is identifying users who are likely to become victims of phishing attacks. If a phishing prevention system could find that the user would disclose their personal information, there might be a chance to protect them from phishing. In order to develop such systems, it is necessary to understand what the users are really thinking. In the context of cognitive psychology, such internal mental processes can be estimated by observable and measurable behavior. Thus, we employ the analysis of the eye movement patterns for phishing prevention.

To achieve this goal, we explore a suitable method for recognizing how users do to assess the credibility of websites. We consider such case in which a user assesses a website by its contents, and ignore meaningful signals displayed in the browser's address bar. Figures 1 and 2 respectively show the heat maps of the eye fixation locations and durations on both phish and legitimate website for novice users (a) and expert users (b). The color red denotes the areas that attracted the user's gaze the most and green denotes moderate gaze activity. In the phishing case, the novice looked at the web content but ignored the browser's address bar while assessing credibility, as shown in Figure 1a. Since the text and visuals in phishing sites are quite similar to the ones in legitimate sites, the novice failed to label the phishing site correctly. In the legitimate case, the novice also only paid attention to the content of a web page as shown in Figure 2a. By contrast, an expert tends to evaluate the site's URL and/or the browser's SSL indicator rather than the contents of the web page to judge the credibility of the sites, as shown in Figures 1b and 2b.

We therefore hypothesize that the analysis of eye movement on the particular areas of interest (AoIs) would allow to extract what are the criteria that helped the user in making a trust decision. To assess our hypothesis, we conducted two types of participant-based experiments. In the first experiment, we analyze

Table 1: Conditions of each site used for recording eye movement

#	Website	Phish	Lang	Description
1	Google	no	JP	SSL
2	Amazon	yes	JP	`tigratami.com.br`, once reported as a compromised host
3	Sumishin Net Bank	no	JP	EV-SSL
4	Yahoo	yes	JP	`kazuki-j.com`, once reported as a compromised host
5	Square Enix	yes	JP	`secure.square-enlix.com`, similar to legitimate URL `secure.square-enix.com`
6	Ameba	no	JP	non-SSL
7	Tokyo Mitsubishi UFJ Bank	yes	JP	`bk.mufg.jp.iki.cn.com`, similar to legitimate URL `bk.mufg.jp`
8	All Nippon Airways	yes	JP	IP address
9	Gree	no	JP	non-SSL
10	eBay	no	EN	EV-SSL
11	Japan Post Holdings	yes	JP	`direct.yucho.org`, SSL
12	Apple	yes	JP	`apple.com.uk.sign.in...`
13	DMM	no	JP	SSL
14	Twitter	yes	JP	`twittelr.com`
15	Facebook	yes	JP	IP address
16	Rakuten Bank	yes	JP	`vrsimulations.com`, once reported as a compromised host
17	Sumitomo Mitsui Card	yes	JP	`www.smcb-card.com`, SSL
18	Jetstar Airways	no	JP	SSL, non pad-lock icon by accessing non-SSL content
19	PayPal	yes	EN	`paypal.com.0.security-c...`
20	Tokyo-Tomin Bank	no	JP	3rd party URL `www2.answer.or.jp`, EV-SSL

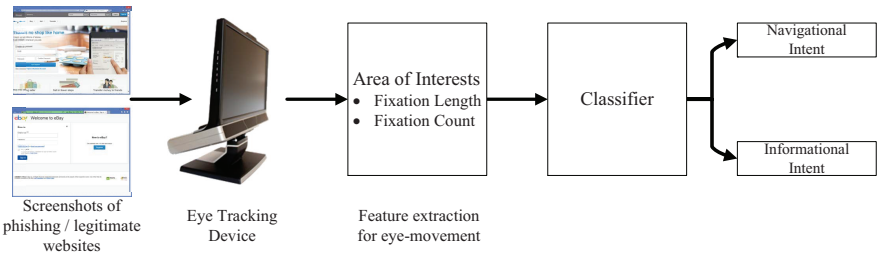


Fig. 3: Block diagram of the experiment

the correlation between eye movements and decision criteria to confirm whether eye fixations can be used as decision criteria indicators. The second experiment investigates whether the eye movement allow to estimate the likeliness of a user to fall victim to phishing.

4 Experiment Setup

This section introduces the procedures of our experiments. Individuals were recruited through a poster advertisement at a college campus during the period of November 2013 - February 2014. Of the 23 participants, a majority of the participants were males in their twenties.

As ethical issue is a concern in our organization, the participants were told they were participating in a security research study. One of the most important ethical rules was that all participants must give their informed consent before taking part in our experiments. The ethical rules also stipulate the need to explain “Why we observe?”, “What we observe?”, “How we observe?” and “Who uses the observed data?”.

(a) Web Content (b) Address Bar

(c) Security Icon for EV-SSL websites (d) Security Icon for SSL websites (e) Surrounding Area of Security Icon

Fig. 4: Definition of AOI

Our prepared twelve phishing sites and eight legitimate sites are listed in Table 1. It should be noted that what we call "phishing sites" were not real phishing sites in the wild, in order to avoid any participant's information leakage. Instead, our participants were presented with the screenshots of a browser that rendered the websites. These screenshots have been taken on Windows 7 equipped with IE 10.0.

Figure 3 shows the block diagram for recognizing the participants' intention. In our experiments, we employed a Tobii TX300 eye tracking system to analyze the eye movement data. With their consent, we measure their eye movements after we calibrated the eye tracking device for each participant. The participants were also shown several options to indicate their decision's criteria: "Content of Web page," "URL of the site," "Security Information of Browser," and "Other Reason." The participants were requested to mark all options that applied (multiple answers allowed), and described in details their reason when selecting the "Other Reason" option.

5 Eye Movement Analysis

5.1 Extraction of implicit intention

We hereafter examine the feasibility of extracting implicit intention from observing the user's eye movements. Based on the number and duration of fixations in each Area of Interest (AOI) of a given input stimulus image, we construct classifiers with SVM to differentiate the participant's implicit intention into navigational and informational intentions.

At first, we evaluate such hypothesis that *the analysis of the eye movement can extract a user's intentions while watching web pages*. Since novices tend to assess credibility by the "Content of Web page," their eye movements would be different from the eye movements of experts. The feature vectors include the number and duration of fixations towards the web content AoI (as shown in

Table 2: Participants' recognition performance by eye movement analysis

Tyep of Intent		AER	AUC
Content of Web page			
	entire time period	32.4%	0.741
	initial ten seconds period	32.2%	0.759
URL of the site			
	entire time perod	28.0%	0.741
	initial ten seconds period	27.8%	0.759
(removed noise)			
	entire time perod	21.3%	0.890
	initial ten seconds period	19.7%	0.917
Security Information from browser (AOI of the address bar)			
	entire time epriod	14.5%	0.855
	initial ten seconds period	14.3%	0.855
(AOI of the padlock icon)			
	entire time epriod	13.5%	0.841
	initial ten seconds period	13.7%	0.809

Figure 4a). The objective variable is a binomial value that denotes whether the participant checked the "Content of Web page" option or not in our questionnaire. The average error rate (AER) was 32.4% and the area under the curve (AUC) was 0.741 as shown in Table 2. Additionally, we assumed that some participants would try to find some trustworthiness information as soon as they have begun browsing the websites. From this perspective, we also extracted the fixation count and duration within the first ten seconds. In this case, the AER was 32.2% and the AUC was 0.759. Hereafter, "initial ten seconds period" means the analysis of eye movement within the first ten seconds, and "entire time period" means the analysis of the entire time while making decision.

We wished for the participants to check the browser's address bar intentionally since the browser's attention on address bar gives trustworthy information such as the URL and security related information. The feature vectors are the number and duration of fixations towards the address bar AoI (as shown in Figure 4b), and the objective variable is a binomial value that denotes whether the participant checked the "URL of the site" option or not in our questionnaire. The AER was 28.0% and the AUC was 0.741 in the case of the entire time period. In the experiment, we found that several participants labeled "URL of the site" as their decision making criteria without actually gazing at the address bar. Even when we redefined the AoI in order to add the surrounding margins, as shown by the green rectangle in Figure 4b, their eye fixations towards the AoI still could not be accounted for. If we remove such falsely motivated decisions, the AER would be 21.3% and the AUC would be 0.890. Additionally, the margined AoI did not improve the performance: in the case of the entire time period, the AER was 22.1% and the AUC was 0.842.

We also assumed that some participants would choose to look at the address bar to find a security indicator. The feature vectors are the number and duration of fixations for that particular AoI, and the objective variable is a binomial value that denotes whether the participant checked the "Security information of browser" option or not in our questionnaire. The possible AOIs are the address bar and the padlock icon. The AER was 14.5% and the AUC was 0.855.

Table 3: Estimation of participants who were going to be victims of phishing

Area of Interest		AER	AUC
Web Content			
	entire time period	24.8%	0.799
	initial ten seconds period	25.1%	0.818
Address Bar			
	entire time perod	24.1%	0.820
	initial ten seconds period	25.7%	0.815
Security Icons			
	entire time perod	24.4%	0.782
	initial ten seconds period	24.8%	0.761
All types of AOIs			
	entire time perod	20.7%	0.873
	initial ten seconds period	21.1%	0.853

We defined the AOIs of the padlock icon, as shown in Figure 4c, 4d, and 4e, for an EV-SSL certificate where the AoI is around the name of the entity as well as the padlock icon, for an SSL certification, it is a rectangle around the padlock icon, and in the case of non-SSL websites, the AOI was a surrounding area for icons displayed in the address bar, respectively. For this last case, we assumed that some participants would check the nonexistence of the SSL certificates. In total, the AER was 13.5% and the AUC was 0.841.

We found that some participants tend not to check the "Security Information of Browser" option, even when the website displayed an SSL padlock icon. The predictor therefore indicates that all participants did not intentionally look at this AoI. We concluded that the AoIs were not as useful to construct a good predictor, however, the AER was 7.6% and the AUC was 0.785. In the case of the websites that displayed an EV-SSL padlock icon, the AER was 33.3% and the AUC was 0.711. When the websites had no certificate, the AER was 10.5% and the AUC was 0.775.

5.2 Estimation of participant's likelihood to be victim

In this experiment, we hypothesized that *the analysis of the eye movement can estimate whether or not a user is going to fall victim to phishing.* The feature vectors in this scenario are the number and duration of fixations towards the three types of AOIs (web content, address bar, and security icons). The objective variable is a binomial value that denotes whether the participant judged correctly or not.

The results are shown in Table 3. By using the combination of all types of AoIs, we observed that the AER was 20.7% and the AUC was 0.873, in the case of the entire time period. The lowest error rate was observed at 8.7% in Websites 10 and 15, and followed by Websites 1, 4, 8, and 9 with 13.0%. Since Website 10 is displayed in English, and since a significant number of the participants were non-native English speakers, we therefore assume that the participants had attempted to assess the website based on the address bar rather than the content.

Additionally, we performed a 10-fold cross validation with tuning parameters by grid search. The results showed that the AER was 29.3% in the case of the entire time period, and 30.8% in the case of the initial ten seconds period.

6 Follow-up Study

In order to thwart bias, we conducted a follow-up study to our experiments for another set of users The study was conducted in September 2014 with 33 new participants. Of the 33, three were female and the rest were male. Twenty of the participants were in their twenties, eight in their thirties, four in their forties, and one in their teens. All of them were attendees of a domestic workshop held in Japan, and were mainly network researchers. The participants were volunteers, and the experiments were done in a conference hall. The rest of the experimental conditions were the same as in the previous experiments.

The analysis of the follow-up study also found that there is a correlation between their eye movement patterns with regards to the AoIs and the criteria they indicated while assessing the credibility of the websites. In the case of web content, the AER of the predictor was 26.7% and the AUC was 0.787.

We then analyzed the eye movements with respects to the AoI of the address bar and their assessment based on "URL of the site". The performance for the intention of checking the URL was observed with the AER of 18.7% while the AUC was 0.837. We also measured the performance for the intention of gaining security information, and found that the pair of the AER and AUC was (18.0% and 0.806) in the case of the address bar, (17.7% and 0.779) in the case of the padlock icon.

We finally observed the feasibility of predicting the participant's likelihood to fall victim of phishing. By using the combination of all types of AoI, the AER was 15.2%, and our ten-fold cross validation raised it to 21.5%.

7 Conclusion

In this paper, we presented a user study in which we evaluated the correlation between eye movements and phishing identification. We used both the duration and the number of eye fixations with respects to particular AoIs, including the area of the rendered web content, the address bar, the padlock icons and their surrounding area. We categorized eye movement patterns along with the types of human implicit intentions.

We conducted a set of experiments which focused on verifying if the eye movement analysis was able to extract users' intentions for assessing the websites' credibility, and to estimate users who were likely to fall victim to phishing. Our result showed that the average error was 32.4% if users assessed the credibility of the website by paying attention to web content, 21.3% if users looked at the URL of the site, and 13.5% if users checked the security information of the browser. We also verified our ability to predict the likelihood that users may fall victim to phishing attacks on the basis of the analysis of their eye movement patterns, and found that it can be estimated with a probability of 79.3%.

Although there still remain other factors in decision-making which must be investigated, we believe that this paper proposed a novel prediction methodology for phishing identification and demonstrates its feasibility. We hope that this

work can help utilize insights from cognitive psychology in order to help protect people from cyber threats.

Acknowledgment

This research has been supported by the Strategic International Collaborative R&D Promotion Project of the Ministry of Internal Affairs and Communication, Japan, and by the European Union Seventh Framework Programme (FP7/2007-2013) under grant agreement No. 608533 (NECOMA). The opinions expressed in this paper are those of the authors and do not necessarily reflect the views of the Ministry of Internal Affairs and Communications, Japan, or of the European Commission.

References

1. Abad, C.: The economy of phishing: A survey of the operations of the phishing market. First Monday 10(9) (2005)
2. Grootjen, M., Neerincx, M.A., van Weert, J.C.: Task Based Interpretation of Operator State Information for Adaptive Support. Tech. rep., ACI/HFES-2006 (2006)
3. Haag, A., Goronzy, S., Schaich, P., Williams, J.: Emotion Recognition Using Bio-Sensors: First Steps Towards an Automatic System. In: Proceedings of Affective Dialogue Systems, Tutorial and Research Workshop. pp. 36–48 (June 2004)
4. Jang, Y.M., Mallipeddi, R., Lee, S., Kwak, H.W., Lee, M.: Human intention recognition based on eyeball movement pattern and pupil size variation. Neurocomputing 128, 421–432 (2014)
5. van Kuilenburg, H., Wiering, M., den Uyl, M.: A Model Based Method for Automatic Facial Expression Recognition. In: Proceedings of the 16th European Conference on Machine Learning (Oct 2005)
6. Leigh, R.J., Zee, D.S.: The Neurology of Eye Movements. Oxford University Press, 4th edn. (1991)
7. Park, U., Mallipeddi, R., Lee, M.: Human Implicit Intent Discrimination Using EEG and Eye Movement. In: Proceedings of the 21st International Conference on Neural Information Processing (Nov 2014)
8. Premacka, D., Woodruffa, G.: Does the chimpanzee have a theory of mind? Behavioral and Brain Sciences 1, 515–526 (1978)
9. Slaghuis, W.L., Holthouse, T., Hawkes, A., Bruno, R.: Eye movement and visual motion perception in schizophrenia I: Apparent motion evoked smooth pursuit eye movement reveals a hidden dysfunction in smooth pursuit eye movement in schizophrenia. Experimental Brain Research pp. 399–413 (2007)
10. Tokuda, S., Obinata, G., Palmer, E., Chaparro, A.: Estimation of mental workload using saccadic eye movements in a free-viewing task. Proceedings of the 33rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society pp. 4523–4529 (August 2011)
11. Watson, D., Friend, R.: Measurement of social-evaluative anxiety. Consulting and Clinical Psychology 33, 448–457 (1969)