Stimuli for Gaze Based Intrusion Detection

Ralf Biedert[1], Mario Frank[1], Ivan Martinovic[2], Dawn Song[1]

[1] University of California, Berkeley
[2] University of Oxford

Abstract. User authentication is an important and usually final barrier to detect and prevent illicit access. Nonetheless it can be broken or tricked, leaving the system and its data vulnerable to abuse. In this paper we consider how eye tracking can enable the system to hypothesize if the user is familiar with the system he operates, or if he is an unfamiliar intruder. Based on an eye tracking experiment conducted with 12 users and various stimuli, we investigate which conditions and measures are most suited for such an intrusion detection. We model the user's gaze behavior as a selector for information flow via the relative conditional gaze entropy. We conclude that this feature provides the most discriminative results with static and repetitive stimuli.

1 Introduction

A common problem in security is the detection and prevention of illicit access. There are a multitude of security measures in use to prevent illicit access, e.g., PINs, tokens or asymmetric keys [9]. However, once 'authenticated' no subsequent checks are performed, rendering the system vulnerable to various attacks in case the user forgets or is prevented from logging off—or simply if his token is being stolen or copied. A possible solution is to replace single point authentication with a continuous biometric one [1] [13] relying on features inherent to the individual itself, such as the iris [3], facial features [2], or fingerprints. However a problem with these methods is that they can, sometimes by the most simple means [12], be copied or forged likewise. It has also been suggested to use eye tracking as an interactive *biometrics* based on shared secrets (e.g., [5], [4] or [10]) and task dependent user behavior [7] [6], which could render some of the attacks impossible or at least harder. However, most of the presented methods still require explicit interaction.

With these observations in mind, we reconsider the use of task learning effects [6] for the *detection* of illicit access. The key idea is that a user in her usually highly individualized working environment is the only person inherently familiar with the layout and content of many of the frequently presented stimuli. Other users are likely to produce a significantly higher amount of searching, reading or comprehension behavior in situations to which the true user is already accustomed to. Eventually this searching and sense-making behavior is likely to reflect in the user's eye movements and there is a chance for the system to observe it through eye tracking.

Obviously, not every task or stimulus will be equally suited to detect familiarity. Hence, in this paper we will investigate three different types of stimuli

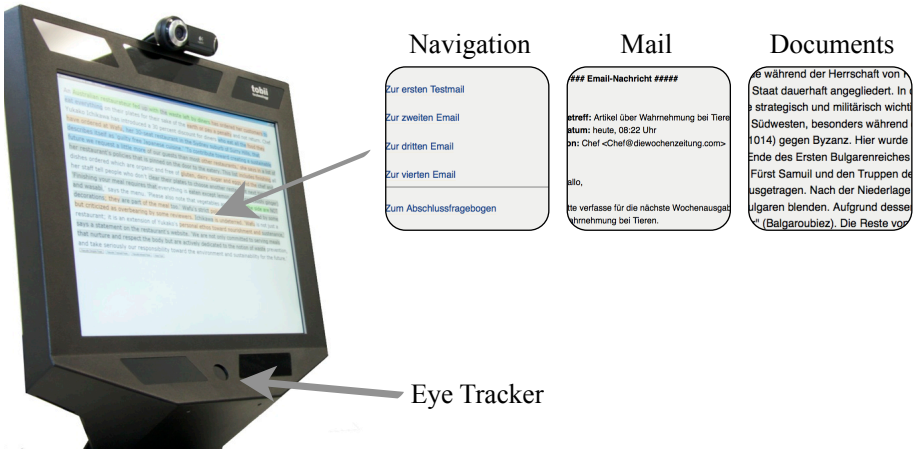

Fig. 1. Outline of the proposed system: the users will encounter various well known points in their system interaction, such as document overviews or dialog text. By observing their gaze behavior and interaction on these stimuli we try to estimate to what extent the user is familiar with the presented information. Unfamiliarity provides evidence of an ongoing intrusion.

and two metrics. We also report how candidates for a more detailed subsequent analysis should look like. In Section 2 we outline our theoretical considerations on the proposed detection system. In Section 3 we present an eye tracking experiment in which we recorded the interaction of a group of 12 users on three different stimuli types. In Section 4 we provide an analysis of various features to distinguish familiar from unfamiliar users. We conclude with Section 5, outlining how we will continue our research.

2 Working Hypothesis

We base our intrusion detector on task learning effects [6]. The learning effect assumes that users become increasingly familiar with certain tasks and layouts converging their behavior to more *optimized* or *informed* pattern. In the daily interaction, a legit user would operate the system normally and encounter various key points, most of which he should rapidly pass with little effort. In contrast, an attacker would be challenged with the additional burden of first comprehending the interface before he would be able to interact with it. Over time the system can therefore compute a hypothesis on the user's familiarity with the presented responses. Should the user's reaction after one or several passes accumulate evidence that he is not familiar with the system this is an indicator for possible compromising.

3 Experiment

We prepared a workflow in which participants are asked perform three tasks, simulating how a user would interact and familiarize with a working environment. They check for mails in a web based interface and receive messages from their hypothetical supervisor, instructing them to read and research on various topics. The experiment consists of four rounds. Each round is structured in the same way. It starts with an overview page that reminds the users of their task. This page also contains navigational links to various mails. Each mail is written by the users imaginary employer, contains a visible mail header, the same introduction and a variable part describing their task for this round. It also contains links to two or three attached documents about different encyclopedic topics. In the end, the users are presented a questionnaire that contains various comprehension questions to ensure they really do as being told.

We invited 12 users (11 of them male), their average age was 24.6 years and most of them were university students. We did not reveal the true purpose of the study, instead users were told to participate in a general text comprehension study. The overall time for the experiment ranged between 30 to 45 minutes. For the whole interaction session, the user's gaze data, along with a number of other features, was recorded on a Tobii 1750 device, sampled with 50 Hz (compare Figure 1).

4 Evaluation

Our analysis of the recorded data focuses on two aspects. We start by analyzing and comparing a number of metrics to model familiarity on a stimulus when considering eye movements, with a special focus on the relative entropy of perceived information. In the second part of the evaluation we then train a classifier on a subset of users and predict how such an algorithm would generalize.

4.1 Features for detecting users in an unfamiliar environment

As stated, the working hypothesis is that users who are unfamiliar with the working environment can passively be discriminated from users that are familiar with this environment. To test this hypothesis we develop a set of features that can be extracted while a user is working at a computer screen. In the following we describe these features and in Section 4.2 we report on our experimental findings of a user study. For the remainder of this paper we consider especially fixations. All recorded data has been de-noised with an independent median filter and fixations are detected using a ($25px$, $100ms$) dispersion window.

Processing time Probably the simplest feature that one can use to discriminate informed users from uninformed users is the processing time for a task. In practice, such a task can have several forms. For instance it could be the task of selecting an icon in a window containing a set of icons. Here processing starts as soon as the user opens the window document, the end point is be the first click on a navigation link. However, since document or stimulus processing can often be interrupted we only focus on the time the user was actually considering the page, this is: the sum of all fixation times on that stimulus.

Relative conditional gaze entropy We assume that an uninformed user would scan the stimulus for relevant information while the informed user would find the relevant information more directly[11]. If this assumption holds then the gaze of an uninformed user would visit the set of objects with potential relevance in an unpredictable way while for the informed user it would be easy to predict that she gazes at the relevant objects.

To quantitatively measure this information we propose the relative conditional gaze entropy (RCGE) as follows. Let $\mathcal{O} = \{o_1, o_2, ..., o_n\}$ be the areas of the screen that show candidate objects (for instance, the set of icons in a window or the set of words in a text). One such area is the set of pixels showing one object. We treat a gaze point as a random variable G and consider g to be a random outcome of a trial, effectively modeling the stimulus as an information source operated and observed by the user. Then the relative conditional gaze entropy (RCGE) is

$$H(G|g \in \mathcal{O}) := -\log_2(n)^{-1} \sum_{i=1}^{n} p(g \in o_i | g \in \mathcal{O}) \log_2(p(g \in o_i | g \in \mathcal{O})) \qquad (1)$$

Here $g \in o_i$ means that the user fixates at one particular object o_i and $g \in \mathcal{O}$ means that the user's gaze g is located in anyone of the candidate areas. By dividing with the entropy $\log_2(n)$ of the uniform distribution, this measure becomes a relative entropy score between 0 and 1, where 1 refers to maximum entropy and 0 means the user stares at a single object o^*.

Let $\mathcal{G} = \{g_1, g_2, ..., g_k\}$ be the measured gaze coordinates of a user working at the computer screen. We then estimate the involved probabilities $p(g \in o_i | g \in \mathcal{O})$ by the relative object frequencies h_i

$$h_i := \left(\sum_{j=1}^{k} I(g_j \in o_i) \right) \left(\sum_{j=1}^{k} I(g_j \in \mathcal{O}) \right)^{-1} \qquad (2)$$

Here, the indicator function $I(p)$ returns 1 if the predicate p is true and 0 if it is false. Note that we designed this measure in a general way. It can be applied to content of various kinds such as, for instance, documents, icons, options in drop-down lists, or any other information carrying element.

In our experiment we defined a geometric bounding box for each of the words in the document and counted the relative word frequencies h_i while the users processed the document. For tasks other than reading, it might be advantageous to record gaze over a constant time window and count frequencies for other geometric bins.

4.2 Results

We compute the described features for three kinds of documents that the users repeatedly encountered during the experiment. The results can be seen in Figure 2. From left to right the results refer to a constant description of some rules (this is always the same document), the emails from a fictional supervisor explaining the current task (parts of this mail change over the repetitions), and the working documents themselves (completely varying content).

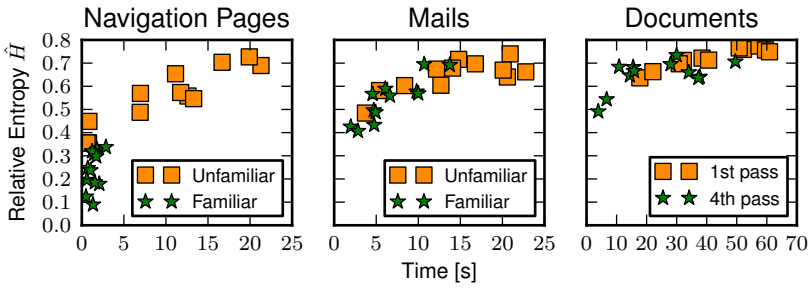

Fig. 2. Relative conditional gaze entropy versus processing time of unfamiliar users (users who see this kind of document for the first time) and informed users (who see it the fourth time). Measurements are carried out on three different kinds of documents: Constant navigational content (navigation pages, left) and varying information in a particular part of the text (mails, middle position). Note that for the documents (right) the content of the first and forth pass differ. They serve as a baseline comparison for our features.

We observe that while a user repeatedly works with documents of one kind both the processing time and the RCGE decrease, as can be seen in Figure 2, left. The class of familiar users and the class of new users are linearly separable in this 2D feature space. The less relevant the content of a document is the stronger is this effect. For classes of documents with lots of varying relevant content, the informed user can not be discriminated from the uninformed user, that is: their RCGE scores and processing times are not separable (see for instance Figure 2, right). This influence of content relevance has an impact to the choice of screen content that can be used for intrusion detection. The detector requires screen content that is ideally constant over many repeated visits or that changes only little (such as documents that change in content only at a few places). This insight guides our future work towards using stimuli as, for instance, desktop icons or menu items.

As a second observation we note that processing time is less reliable than RCGE. Some of the users processed the constant rule screen at first sight as fast as users that are familiar with the screen. Obviously, it is possible to realize quickly that a screen content is not relevant at all. However, this still requires some search over the different objects of interest such that the gaze pattern of these unfamiliar but fast users is more entropic than those of familiar users.

5 Conclusion and Outlook

We presented a novel idea for assisting intrusion detection with eye tracking data. We showed that modeling textual key points as an information source, *transmitted* to the eye for task processing, we can distinguish informed users from uninformed users for some stimuli and tasks. Based on such an information transmission model it appears that especially information-rich but static stimuli which are encountered repeatedly provide the best grounds for such a detection.

There are a number of interesting directions we want to address in the future. We especially consider file- and folder navigation to be highly promising, since these structures are relatively stable over time and are frequently visited while working on a machine. Another possible direction might be the investigation of actual behavior on individual *key words*. Since a familiarity effect on words and its influence on fixation time (compare [8]) is well known, the observation of fixation times on specific low frequency words that are familiar to the user could also be an option to investigate.

References

1. Awad, A., Ahmed, E.: Detecting Computer Intrusions Using Behavioral Biometrics. Department of Electrical and Computer Engineering, University of Victoria (2005), http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.66.2289
2. Bledsoe, W.: The model method in facial recognition, Panoramic Research Inc. Palo Alto (1964)
3. Daugman, J.G.: High confidence visual recognition of persons by a test of statistical independence. IEEE Transactions on Pattern Analysis and Machine Intelligence 15(11), 1148–1161 (1993), http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=244676
4. De Luca, A., Weiss, R., Hußmann, H., An, X.: Eyepass - eye-stroke authentication for public terminals. CHI EA '08: CHI '08 extended abstracts on Human factors in computing systems (Apr 2008), http://portal.acm.org/citation.cfm?id=1358628.1358798
5. Dunphy, P., Fitch, A.: Gaze-contingent passwords at the ATM. The 4th Conference on Communication by Gaze Interaction – Communication, Environment and Mobility Control by Gaze (2008)
6. Kasprowski, P.: Human identification using eye movements. Doctoral thesis (2004), http://www.kasprowski.pl/phd/
7. Kinnunen, T., Sedlak, F., Bednarik, R.: Towards task-independent person authentication using eye movement signals. In: ETRA '10: Proceedings of the 2010 Symposium on Eye-Tracking Research & Applications. School of ComputingUniversity of Eastern Finland, Joensuu, Finland, ACM Request Permissions (Mar 2010), http://dl.acm.org/citation.cfm?id=1743666.1743712
8. Reichle, E.D., Rayner, K., Pollatsek, A.: Eye movement control in reading: accounting for initial fixation locations and refixations within the E-Z Reader model. Vision Research 39(26), 4403?4411 (1999)
9. Schneier, B.: Secrets and Lies - Digital Security in a Networked World. Wiley (2004), http://www.worldcat.org/title/secrets-and-lies-digital-security-in-a-networked-world-with-new-information-about-post-911-security/oclc/443509536
10. Weaver, J., Mock, K., Hoanca, B.: Gaze-based password authentication through automatic clustering of gaze points. In: Systems, Man, and Cybernetics (SMC), 2011 IEEE International Conference on. pp. 2749–2754 (2011),
11. Wyatte, D., Busey, T.: Low and high level changes in eye gaze behavior as a result of expertise. Eye 60(70), 80–90 (2008)
12. Xiao, Q.: Security issues in biometric authentication. In: Information Assurance Workshop, 2005. IAW '05. Proceedings from the Sixth Annual IEEE SMC. pp. 8–13 (2005),
13. Yan, J.: Continuous authentication based on computer security. Tech. rep., Luleå University of Technology (2009), http://epubl.ltu.se/1653-0187/2009/005/LTU-PB-EX-09005-SE.pdf