Platform for Evaluation of Readers' Implicit Feedback using Eye-Tracking

Miroslav Živković
Department of Information and Computing Sciences
Utrecht University
Utrecht, The Netherlands
m.zivkovic@uu.nl

Egon L. van den Broek
Department of Information and Computing Sciences
Utrecht University
Utrecht, The Netherlands
vandenbroek@acm.org

Frans van der Sluis
Department of Information Studies
University of Copenhagen
Copenhagen, Denmark
f.vandersluis@acm.org

ABSTRACT

Large amounts of information are nowadays easily obtainable using the Internet, and using implicit feedback whether a reader finds a text interesting is desirable. Eye-tracking technology could be used for such a feedback, and a combination of eye-movement features and a textual complexity measure can be used to predict the user's interest. In this paper we give an overview of a platform developed to evaluate and visualize implicit feedback of a person who reads a text. Based on the eye-movement samples provided, a model is trained that could be used to predict comprehensibility of a user reading a text. This prediction is combined with objective complexity evaluation of the text using data mining methods, and the outcome is used to select a text (from a repository) that a user may find more valuable (interesting). We briefly discuss the requirements, architecture and implementation of this platform.

CCS CONCEPTS

• **Information systems** → *Information retrieval*; • **Human-centered computing** → *Human computer interaction (HCI)*; • **Software and its engineering** → *Software architectures*;

KEYWORDS

Information experience, eye-tracking, Java, WEKA.

ACM Reference Format:
Miroslav Živković, Egon L. van den Broek, and Frans van der Sluis. 2018. Platform for Evaluation of Readers' Implicit Feedback using Eye-Tracking. In *ECCE'18: Proceedings of the 36th European Conference on Cognitive Ergonomics (ECCE2018), September 5–7, 2018, Utrecht, Netherlands.* ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/3232078.3232099

ECCE'18, September 5–7, 2018, Utrecht, Netherlands

ACM ISBN 978-1-4503-6449-2/18/09...$15.00
https://doi.org/10.1145/3232078.3232099

1 INTRODUCTION

Understanding and knowing what the user wants in terms of information is a paramount challenge for information systems [7, 9]. Asking users to provide copious and continuous input about the information they want is not likely to succeed; instead, implicit feedback that does not require any interaction from the user is a more viable option. A particularly promising opportunity for implicit feedback on information value comes from eye-tracking data [12]. Devices such as Google Glass, eye-tracking fitted tablets, and digital watches not only create new platforms to deliver information; but, also lead to a surge of available behavioral and physiological data for implicit feedback on information value. However, the challenge is not only to analyse the new signals (e.g., physiological data); but, also to match signal characteristic with textual features.

How could one determine what information is interesting to the user, or, more specifically, evokes the interest emotion? According to Silvia [8], interest is invoked by two appraisals of an event. The first appraisal evaluates the novelty-complexity of an event, and the second appraisal evaluates the comprehensibility of an event. Van der Sluis et al. [11] verified that this theory also applies to text: they define a sweet spot of interest where textual complexity is high enough to make for a challenging text, but not so high that the text becomes incomprehensible. The objective textual complexity (derived from the text by a model, without human input) and appraised comprehensibility can be used to predict and explain interest. Such a system could derive the objective textual complexity from the text using a model (or, an alternative possibility is to use the mean or median of complexity ratings given by readers), and it is known that eye movement data gathered by an eye tracker can be used to assess comprehension processes [5].

While eye-tracking seems to offer a lot of potential as implicit feedback mechanism, there is a lack of standardized software for analysis, visualization and processing of eye-tracking signals. There is a need for a platform that would integrate eye-tracking signal processing and text comprehension inference from such a signal. As we want to predict user interest in offered text, our platform must support complexity calculations as well, based on text mining.

In this paper we present an ongoing effort to build a platform that

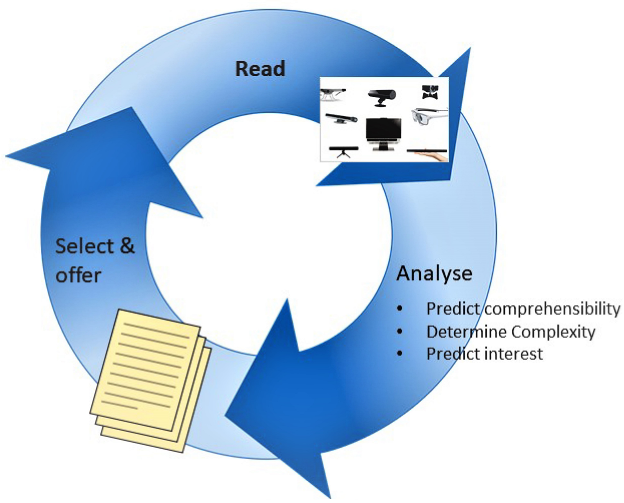

Figure 1: Usage scenario: A closed-loop model [10] that uses eye-tracking data while a user reads a document, processes this data, and uses it to predict text comprehensibility.

- integrates different mechanisms for eye-tracking signal processing which could be fine-tuned, added or removed
- uses eye movements as implicit feedback on comprehensibility when user reads a text
- integrates text mining approaches to determine text complexity
- infers user interest in read text from complexity and comprehensibility.

This paper is organized as the following: next, we describe the use case scenario which is used to derive the requirements. In Section 3 we describe the architecture of eye-tracking capabilities that need to be integrated. Based on this architecture we have developed our tool that could be used to offer texts a user would find interesting based on eye-tracking data and text analysis. We conclude the paper and discuss the future work in Section 4.

2 USE CASE AND REQUIREMENTS

In a typical use case scenario, illustrated in Figure 1, a user reads a piece of text, while his eye movements are recorded using a commercially off-the-shelf eye-tracking device. Based on eye-tracking data, and using a platform model, comprehensibility is predicted. On the other hand, complexity is calculated, using text mining methods. Based on comprehensibility and complexity, interest is derived [11]. This is used to select a document from repository that may be of more interest to a user. The above-mentioned platform model is used to predict text comprehensibility based on eye-tracking data. This model has to be tuned, and eye-movement features used in this model need to be selected and optimized during an *initial* phase. The tuned model would be used in *exploitation* phase. Our platform should support both phases.

Further, a repository that contains texts offered to users may contain the objective characteristics of texts, such as calculated complexity and predicted interest. During the initial phase the texts' complexity values should be established and stored in a repository.

State-of-the-art software for analysis of the eye-tracking data is probably not as advanced as the research front. Many researchers program their own software, either because the current software is not ideal for their system, or lacks implementation they need. Besides, eye-tracking signals are considered to be noisy, and due to limitations of the signal acquisition, signal reconstruction, de-noising and analysis may be required. In some cases, not all steps may be necessary, and in some other cases, tuning of the parameters for each step is preferred. The majority of the software that is available to the researchers do not offer capabilities as signal reconstruction or de-noising, and completely rely on what is provided by the eye-tracking software delivered with a device. There is a need to provide these signal processing solutions, as well as flexibility to use them.

The requirements could be specified as the following - the platform

- supports both initial and exploitation phase;
- allows that data signals, in different stages of processing could be stored in a database;
- allows that different eye-tracking processing steps may be included or omitted;
- supports different models and classification algorithms for comprehensibility prediction from eye-tracking data. The models are trained/tuned during the initial phase; and
- supports text complexity calculations.

These requirements are used as the input for the architecture we specify in the next Section.

3 ARCHITECTURE

The architecture and intended integration with consumer market eye trackers is illustrated in Figure 2. The architecture supports both the *initial* phase and the *exploitation* usage of our tool. During the initial phase, eye-tracking data are obtained while users read a set of pre-selected documents. Next to this, the users explicitly provide feedback on interest, comprehensibility and complexity of these documents.
There are three essential steps in this phase:

- Text mining is applied to calculate complexity of the documents from the texts repository. State-of-the-art solutions achieve a classification accuracy of up to 93.62% and a correlation with human ratings of up to r = .703 [11].
- Raw eye-tracking data is stored in the database, and then pre-processed (or: "cleaned"). Hereafter, the eye-movement features are extracted (e.g., fixations and saccades) and parameters of these features are calculated. The outcome of this step is aggregated data and is stored in a database.

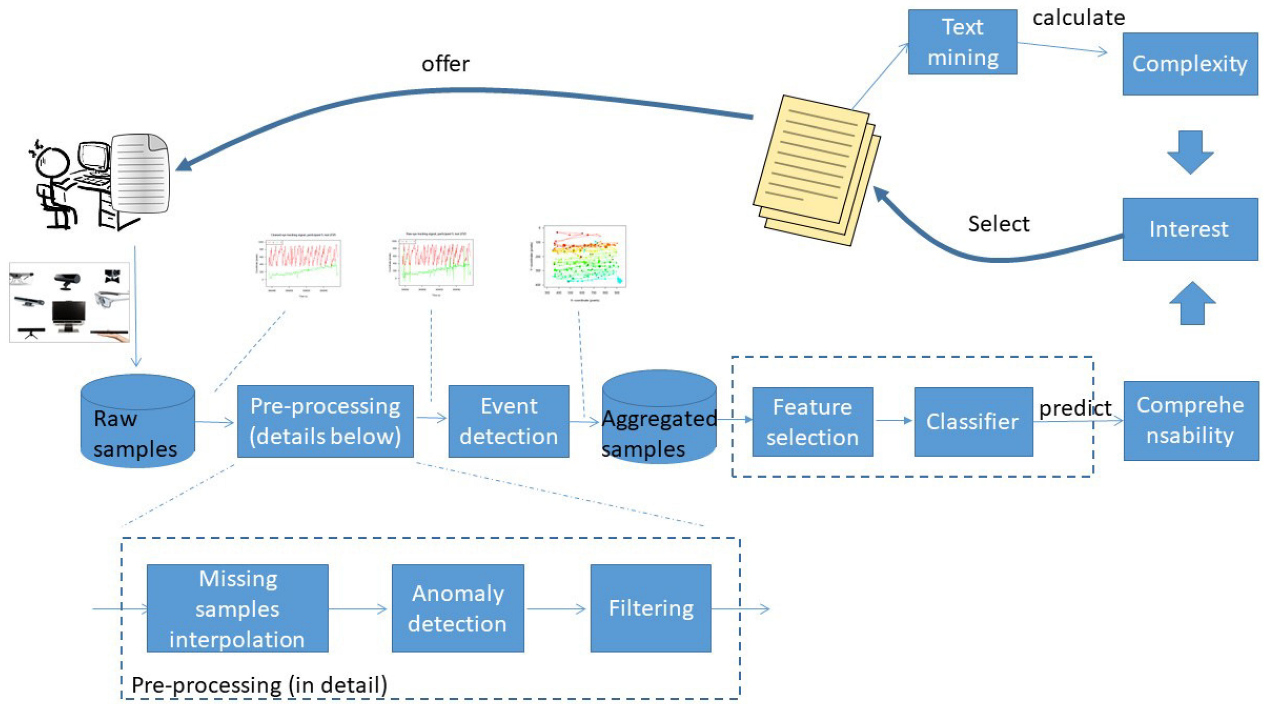

Figure 2: Architecture of the platform for evaluation of readers' implicit feedback using eye-tracking. The raw eye-tracking data is pre-processed (as detailed) before detection of eye movements (events). Visualization is possible after each step. The aggregated data is used in subsequent steps, including feature selection, classification, and prediction [10].

- The aggregated data is used to build the data model and classifier of comprehensibility and interest from the eye-tracking data. Feature selection and reduction is important part of this step, as it mitigates the curse of dimensionality, removes redundancy among the signal's features and their parameters, and, hence, becomes more robust and generic.

The architecture allows for stored raw eye-tracking data to be read from the database, and pre-processed again. This allows to re-create the aggregated data, and consequently, a new classifier may be trained. Therefore our platform supports different algorithms/solutions that could be applied for data processing. In a typical pre-processing scenario, illustrated in Figure 2, the eye tracking signal is first analysed for missing samples. We noticed that eye trackers do not provide samples for all sampling instances; hence, there could be some missing data. In order to overcome this, we interpolate the signal. Next, the signal is fed to the signal cleaner. The cleaner first removes anomalies from the data. When a reading session contains too many anomalies, the entire session is considered to be invalid and is discarded. An example anomaly in the eye tracking data is when gaze position coordinates are negative. This corresponds to tracking loss of the equipment, most likely because of the blinks. We therefore remove these data samples [1]. Similarly, we consider an anomaly for the entire reading session when the number of fixations or saccades is too low. Once the anomalies are removed, we apply a median filter to the eye movement signal to remove noisy spikes observed in the data.

The cleaned signal is then fed to the event detector which employs an algorithm (e.g., the I-DT algorithm [6]) to detect fixations and saccades in the signal. This results in aggregated data, which can be visualized.

The feature extractor uses the cleaned signal and detected events to determine a set of feature values for each reading session (e.g. the number of fixations per document). Each valid reading session is now represented as a feature vector, that are used to train prediction models.

Different prediction models could be deployed in our platform. For example, one could predict comprehension ratings from our extracted features. For interest prediction, a wide variety of approaches was investigated. The authors of [11] show that objective textual complexity (derived from the text by application of text mining, without human input) and appraised comprehensibility can be used to predict and explain interest. A measure for textual complexity could be either the mean or the median of the ratings. interest is predicted, see Figure 3. During the exploitation phase a user

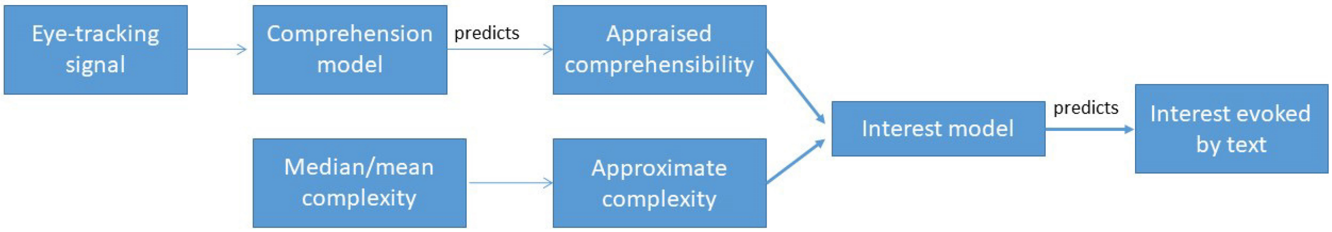

Figure 3: An overview of an interest prediction model. The comprehension model predicts a document comprehension rating from the eye-tracking data, which is further used in combination with the median or mean document complexity rating to predict a user interest.

is offered a document, and eye-tracking data is acquired. The processing flow of this raw data is then identical to the one that is used during initial phase. The aggregated data is used to predict comprehensibility. The resulting, predicted values are used together with calculated documents' complexity from a repository and the user interest is inferred from these. In case when inferred interest is below a certain threshold, a document with higher interest value may be offered to the user. The evaluation data, as well as the eye-tracking data, show which document is read in a session, so we can use this to link evaluation and eye-tracking data together with the respective documents.

We implemented our platform using Java programming language [3]. The models, classification and prediction components of the architecture were developed using WEKA [2]. Finally, text mining and processing used to determine text complexity were implemented using Stanford's Natural Language Processing Toolkit [4]. The machine learning algorithms used for prediction could be pre-selected or selected from a list in the graphical user interface (GUI).

4 DISCUSSION

In this paper, we discussed a platform developed to analyse and visualise implicit feedback of a person who reads a text. Based on the eye-movement samples provided, our platform trains a model that could be used to predict information experience of a user reading a text, i.e. interest. This prediction is further used to steer a text selection that a user would find more valuable. We discussed the main architecture components, as well as underlying data acquisition, data cleanup and models used to derive interest from eye-tracking data as well as objective complexity. Our platform offers a unified approach to acquisition, preparation and processing of the eye movements data. It also allows for an easy integration and exploration of different algorithms and parameters.

The next steps would be to extend the tool in such a way that samples from the eye-trackers are processed on-the-fly. This would require to implement a sort of a "wrapper" for eye-trackers that would be used. Another interesting direction would assume that documents offered would be rich text documents, including images, figures, formulae, etc. [9]. This requires research and implementation of the methods to evaluate complexity of these documents, and, probably, investigating and integrating new eye movement features and parameters. Finally, we may consider implementation in other programming languages, such as Python.

ACKNOWLEDGMENTS

This work has been done within the scope of the NWO IPPSI-KIEM project "Adaptive Text Mining: Information in the Eye of the Beholder" (643.000.002).

REFERENCES

[1] A. T. Duchowski. 2007. *Eye Tracking Methodology: Theory and Practice.* Secaucus, NJ, USA: Springer-Verlag New York, Inc.
[2] E. Frank and I. H. Witten. 2016. *The WEKA workbench. Online Appendix for "Data Mining: Practical Machine Learning Tools and Techniques"* (4th ed.). Morgan Kaufmann.
[3] Java [n. d.]. Java Programming Language. ([n. d.]). http://www.oracle.com/technetwork/java/index.html
[4] C. D. Manning, M. Surdeanu, J. Bauer, J. Finkel, S. J. Bethard, and D. McClosky. 2014. The Stanford CoreNLP Natural Language Processing Toolkit. In *Association for Computational Linguistics (ACL) System Demonstrations.* The Association for Computer Linguistics, 55–60. http://www.aclweb.org/anthology/P/P14/P14-5010
[5] K. Rayner, K. H. Chace, T. J. Slattery, and J. Ashby. 2006. Eye Movements as Reflections of Comprehension Processes in Reading. *Scientific Studies of Reading* 10, 3 (2006), 241–255.
[6] D. D. Salvucci and J. H. Goldberg. 2000. Identifying Fixations and Saccades in Eye-tracking Protocols. In *Proceedings of the 2000 Symposium on Eye Tracking Research & Applications (ETRA '00).* New York, NY, USA: ACM, 71–78.
[7] T. Saracevic. 2007. Relevance: A Review of the Literature and a Framework for Thinking on the Notion in Information Science. Part II: Nature and Manifestations of Relevance. *Journal of the Association for Information Science and Technology* 58, 13 (2007), 1915–1933.
[8] P. J. Silvia. 2008. Interest – The curious emotion. *Current Directions in Psychological Science* 17, 1 (2008), 57–60.
[9] E. L. van den Broek. 2005. *Human-Centered Content-Based Image Retrieval.* Ph.D. Dissertation. Nijmegen Institute for Cognition and Information, Radboud University Nijmegen.
[10] E. L. van den Broek. 2011. *Affective Signal Processing (ASP): Unraveling the mystery of emotions.* Ph.D. Dissertation. Human Media Interaction (HMI), Faculty of Electrical Engineering, Mathematics, and Computer Science, University of Twente, Enschede, The Netherlands.
[11] F. van der Sluis, E. L. van den Broek, R. J. Glassey, E. M. A. G. van Dijk, and F. M. G. de Jong. 2014. When Complexity Becomes Interesting. *Journal of the Association for Information Science and Technology* 65, 7 (2014), 1478–1500.
[12] A. Van Drunen, E. L. van den Broek, A. J. Spink, and T. Heffelaar. 2009. Exploring Workload and Attention Measurements with uLog Mouse Data. *Behavior Research Methods* 41, 3 (2009), 868–875.