2019 IEEE International Conference on Edge Computing (EDGE)

# Edge-based and privacy-preserving multi-modal monitoring of student engagement in online learning environments

Davy Preuveneers and Wouter Joosen
*imec-DistriNet-KU Leuven*
*Heverlee, Belgium*
{*firstname.lastname*}*@cs.kuleuven.be*

***Abstract*—With engagement being an early predictor for a student's learning achievements, it is paramount that teachers can observe the behavior of their audience to keep them engaged, for example, with interactive lectures. In order to address this concern, we present an edge-based multi-modal engagement analysis solution for teachers to maintain an engagement overview of their entire audience, including those in distance learning settings. We designed and evaluated an edge-based browser solution for the analysis of different behavior modalities with cross-user aggregation through secure multiparty computation.**

***Keywords*-multi-modal engagement monitoring; edge computing; browser; privacy; online learning**

## I. Introduction

Student engagement is a topic that has sparked a fair amount of interest over the past decade within the e-learning research community as well as with higher education institutions. Previous research [1], [2] has shown that students who are engaged with their studies are more likely to be successful. However, engagement is a complex construct. Multiple theories have been proposed in the literature and compared by Kahu [3]. Engagement can be understood as a mix of behavioral, cognitive and emotional (or affective) factors. Maintaining a continuous awareness of the engagement level of students is not only a challenge in a face-to-face classroom setting, but even a bigger one in remote settings in which students participate in interactive online lectures:

1) First, to capture behavior engagement, the system should process audiovisual, interaction and physiological data of the audience in near-real time.
2) Second, the effectiveness of different behavior and interaction modalities to measure engagement is highly context-dependent.
3) Last but not least, the continuous tracking and centralized analysis of sensitive personal information may invade the privacy of remote students [4]. This remains an issue today[1] and is a growing concern.

To address these three concerns, we present a decentralized multi-modal engagement analysis solution that distributes the student engagement analysis towards the client devices (e.g. laptops of students). Our edge computing software framework analyses different kinds of student behaviors across different online learning contexts.

[1]https://campustechnology.com/articles/2018/05/02/when-learning-analytics-violate-student-privacy.aspx

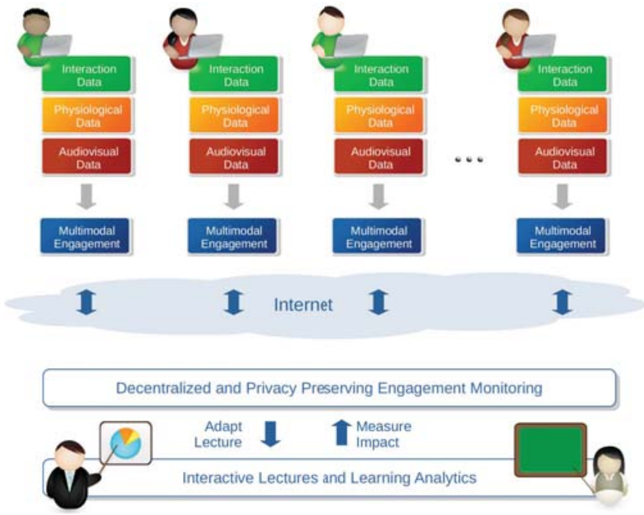


Figure 1. Technology-enhanced multi-modal engagement monitoring of students in blended learning environments

## II. An edge-based framework for privacy-preserving engagement analysis of multi-modal behaviors

To support teachers with retaining the engagement of their students, our engagement analytics framework (as depicted in Figure 1) must be able to analyze in real-time heterogeneous static silos of data (e.g. user profile data of students) and dynamic data streams (e.g. behavioral and physiological measures, click streams, evolving collaborations with peers). Our edge computing solution leverages the extension capabilities of the web browser to analyze student behavior:

1) Keyboard usage
2) Mouse motion events and clickstream data
3) Web camera for head pose estimation
4) Browser tab activations, updates and removals, and website snapshots
5) Browser information and window focus listener

978-1-7281-2708-8/19/$31.00 ©2019 IEEE
DOI 10.1109/EDGE.2019.00017

The main benefits of our browser extension is that it (1) can analyze on- and off-task behavior on the client via the browser of the student across different operating systems and web browsers, (2) reduces the cloud-based workload by shifting computations towards the client, (3) measure the level of engagement in an application or website agnostic manner which simplifies integration with third party resources, and (4) avoid the transmission of privacy-sensitive information to a centralized virtual server for analysis. More details about the browser extension are available at https://people.cs.kuleuven.be/~davy.preuveneers/lecture+/.

### *A. Server-side aggregation of engagement streams with different data processing pipelines*

The browser-based multi-modal engagement analysis is complemented with a server-side backend that aggregates the high-level engagement scores of the individual students. Furthermore, the backend implements the same multi-modal engagement analysis functionality available in the browser. This allows us to systematically compare the performance gains of analyzing the raw data streams on the server versus shifting (part of) the processing workload towards the student's browser.

### *B. Secure cross-user aggregation of engagement scores*

Our edge-based framework further improves the privacy of its users by aggregating the engagement scores of the individual students in a privacy preserving so that a teacher only has an aggregated engagement score for the entire audience. This is achieved by summing the engagement scores through secure multi-party computation. Secure multiparty computation [5] allows computations on encrypted values. This way it is possible to compute an average engagement score without any individual group member revealing her personal engagement score to the others.

## III. Evaluation

This section reports on various qualitative and quantitative evaluation metrics of our solution, identifying performance and impact trade-offs. We measure the overhead of the engagement analysis at the client side, and carry out a systematic performance analysis of centralized versus edge-based engagement analysis on the server – hosting an Intel Core i7-7700 CPU @ 3.60GHz with 32GB of memory – for a growing number of students.

### *A. Performance impact on server for centralized and edge-based engagement analysis configurations*

We collected real-world interaction traces of student behavior during various interactive lectures[2]. From the 25 students that consented to have their data collected, we filtered the top 3 that created the most events, and used those traces to create a baseline for a "worst case scenario" from a performance point of view. These traces were replicated and replayed to simulate from 10 up to 500 concurrent students to analyze the scalability of the back-end of our solution. Figure 2 compares the performance impact in terms of CPU load on the server for both an edge-based and a centralized data processing pipeline. For the centralized variant, we distinguish two variations where the deep learning model for head pose detection [6] is either run on the CPU or delegated to a CUDA card (more specifically, an NVIDIA Titan XP graphics card). It is clear that the edge-based configuration scales to many more users (the maximum capacity of the server was not reached at 500 concurrent students). Furthermore, by sending the raw data to the server in the centralized scenario, the network bandwidth usage is more than 50 times higher due to the need to stream webcam images to the server. The additional overhead is less significant if the webcam video of each student is streamed to all participants anyhow (e.g. via the WebRTC protocol). However, in such a scenario, the bottleneck in terms of concurrent users is the network capacity on the server.

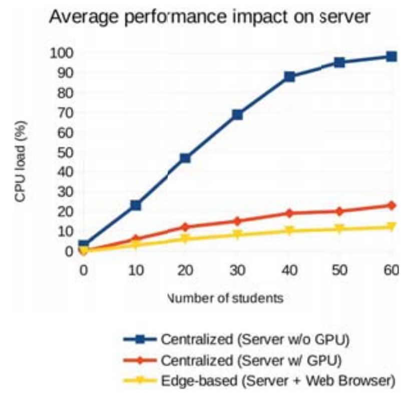


Figure 2. Server-side performance impact in cloud

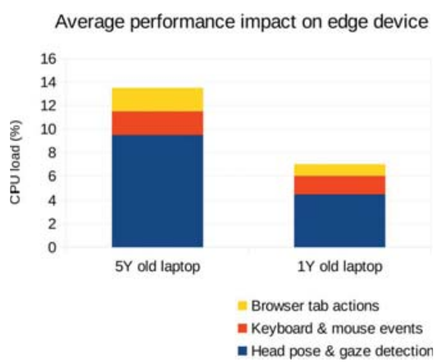


Figure 3. Client-side performance impact (average CPU load over 1 minute) on edge device

[2]The data collection was approved on April 2018 by the Social and Societal Ethics Committee of the university with case number G-2018 04 1206

### *B. Performance impact on edge device for client-side engagement analysis*

Given that the deep learning component for head pose estimation and eye-gaze analysis is computationally the most intensive, we tested the practical feasibility of this component with Google Chrome 69 running on different Dell laptops, the oldest being a Dell Latitude E6330 of more than 5 years old and the newest being a Dell Latitude 7480 of more than 1 year old:

- *5 year old laptop*: Dell Latitude E6330 with Intel Core i5-3320M CPU @ 2.60GHz, 8 GB memory
- *1 year old laptop*: Dell Latitude 7480 with Intel Core i7-7600U CPU @ 2.80GHz, 32 GB memory

For a 640x480 webcam configuration, the maximum framerate including face landmark detection varied between 12 and 22 frames per second, but for engagement analysis one frame per second suffices. Neither of the laptops relied on CUDA hardware acceleration to evaluate the deep learning models. Figure 3 provides a breakdown of the computational overhead of the browser extension into the individual engagement indicators. These values were obtained through the built-in task manager of Google Chrome, which allows to measure the memory footprint, CPU and network usage for each browser tab and extension. It is clear that the JavaScript based Tensorflow model for head pose estimation [6] has the biggest overhead. TensorFlow.js can use a GPU to accelerate math operations. So if available on the device, we expect the performance overhead on the client to be significantly less.

### *C. Privacy impact analysis*

Compared to online learning analytics solutions that centralize and process all raw data in the cloud, our edge-based system provides the following benefits:

- The user remains in control of which data is collected, and must provide consent before any data is captured.
- The individual engagement score is computed on the user's client. No sensitive data (e.g. keystrokes, webcam or website snapshots) are sent to the cloud.
- The cross-user engagement score is aggregated through secure multi-party computation so that individual scores are not revealed to others.

## IV. Conclusion and Future Work

In this work, we presented an edge-based multi-modal engagement solution that runs as an extension with contemporary web browsers. It supports the monitoring of interaction data with the online learning platform as well third party resources, and can analyze head pose estimation using a deep learning model implemented and evaluated in JavaScript. The added value of our contribution is that the multi-modal engagement analysis is off-loaded towards the students' browsers, allowing our framework to easily scale up with a growing number of students. At the same time, it mitigates any privacy concerns that students may have due to continuous tracking of their on-task and off-task behavior. One limitation of our solution is that interaction with non-browser tools (e.g. a native word processing or presentation application) used in the frame of a course are ignored to monitor engagement.

As future work, we will enhance autotuning of the configuration of the browser extension to automatically reduce the performance overhead on the client below a certain threshold. Also, the current solution allows for students to cheat by manipulating the client-side data collection process. As future work, we will investigate the feasibility for temporary replication of the engagement analysis on both the clients and server to detect adversarial behavior.

## Acknowledgment

This research is partially funded by the Research Fund KU Leuven, and by imec through ICON LECTURE+. LECTURE+ is a project realized in collaboration with imec, with Barco, Televic Education and Limecraft as project partners and with project support from VLAIO.

We gratefully acknowledge the support of NVIDIA Corporation with the donation of the Titan XP GPU used for this research.

## References

[1] M. Atherton, M. Shah, J. Vazquez, Z. Griffiths, B. Jackson, and C. Burgess, "Using learning analytics to assess student engagement and academic outcomes in open access enabling programmes," *Open Learning: The Journal of Open, Distance and e-Learning*, vol. 32, no. 2, pp. 119–136, 2017.

[2] E. R. Kahu and K. Nelson, "Student engagement in the educational interface: understanding the mechanisms of student success," *Higher Education Research & Development*, vol. 37, no. 1, pp. 58–71, 2018.

[3] E. R. Kahu, "Framing student engagement in higher education," *Studies in Higher Education*, vol. 38, no. 5, pp. 758–773, 2013.

[4] B. Lorenz, S. Sousa, and V. Tomberg, "Privacy awareness of students and its impact on online learning participation – a case study," in *Open and Social Technologies for Networked Learning*, T. Ley, M. Ruohonen, M. Laanpere, and A. Tatnall, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2013, pp. 189–192.

[5] F. Bayatbabolghani and M. Blanton, "Secure multi-party computation," in *Proceedings of the 2018 ACM SIGSAC Conference on Computer and Communications Security*, ser. CCS '18. New York, NY, USA: ACM, 2018, pp. 2157–2159. [Online]. Available: http://doi.acm.org/10.1145/3243734.3264419

[6] G. Papandreou, T. Zhu, L. Chen, S. Gidaris, J. Tompson, and K. Murphy, "Personlab: Person pose estimation and instance segmentation with a bottom-up, part-based, geometric embedding model," *CoRR*, vol. abs/1803.08225, 2018. [Online]. Available: http://arxiv.org/abs/1803.08225