**Optimizing EEG Visualization Through Remote Data Retrieval**

*N. Capp, C. Campbell, T. Elseify I. Obeid and J. Picone*

The Neural Engineering Data Consortium, Temple University

{noah.capp, tuf46810, tug35668, iobeid, picone}@temple.edu

1. Research reported in this publication was most recently supported by the National Human Genome Research Institute of the National Institutes of Health under award number U01HG008468. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

A common practice in analyzing brain activity is capturing electroencephalograms (EEG), taken by place electrodes on the patient’s scalp which measure electrical activity via voltage differences. These multi-channel voltage signals are commonly stored in the European Data Format (EDF). An EDF file of an unpruned EEG can become quite large – 1MB of an EDF file translates to roughly 1 minute of an EEG recording session with a sampling rate of 250Hz. Additionally, full datasets from the TUH EEG Corpus [1] can reach sizes of over 800GB. Because of the size of the EEG recordings, there is a need for efficient visualization and analysis tools. In this work, we focus on enhancing the user experience by streamlining the process of retrieving EDF data from our server.

An EEG visualization tool [2], written in Python, was developed to facilitate annotation of EEG events in the TUH EEG Corpus [1] and supports many other visualization techniques. Its annotation capabilities are fully integrated within the tool, overlaying the EEG signals; an improvement to software with similar capabilities (e.g. EDF Browser [3] and EEGLab [4]). This software also has spectrogram and signal energy visualizations that are also implemented within the viewing of the raw or filtered EEG signals.

There is cohort retrieval system [5] integrated within the visualization tool that allows search queries to be sent to an EEG dataset that return EDF files relating to the given search terms. This system uses both keyword terms and natural language processing to process the input search query. The medical reports found in the TUH EEG Corpus are then used, along with automatically extracted EEG signal events to rank the most relevant EEG sessions. The tool then allows users to easily visualize and/or annotate EEGs returned by cohort retrieval system.

Previously, visualizing and/or annotating an EEG recording required the EDF file to be locally available on the user’s machine. This often meant our annotators would have roughly 75 GBs of disk space consumed by EDF files. This is less than desirable for a number of reasons involving data integrity and security. An application programming interface (API) has been developed to provide a means of streaming EDF data from our server and into the EEG visualization software. Popular music streaming services provide APIs with similar functionality (e.g. Spotify [6] and SoundCloud [7]).

Through this API, a user can send an HTTP GET Request [8], which contains a file path on the server. The API returns either the paths to the available files in a given directory on the server, or the file itself. This API is protected through a private key, which prevents any unauthorized users from accessing this data [9]. This functionality ensures that users do not need to manually download large batches of EEG data, clearing space on the users’ hard drives and relieving CPU usage on the server. The API is also utilized by the cohort retrieval system [5]. This system returns relative file paths to the most relevant EEG data to the search queries. In previously implementations, this meant the entire TUH EEG Corpus must be locally available on an external hard drive to be able to visualize the data. Instead, the EEG data can simply be streamed into the user’s system using the developed API.

A live demonstration of this integrated visualization and cohort retrieval tool will be provided at the conference. The software is publicly available at *https://www.isip.piconepress.com/projects/tuh\_eeg/ downloads/nedc\_demo*.

References

1. I. Obeid and J. Picone, “The Temple University Hospital EEG Data Corpus,” *Front. Neurosci. Sect. Neural Technol.*, vol. 10, p. 00196, 2016.
2. N. Capp, E. Krome, I. Obeid, and J. Picone, “Rapid Annotation of Seizure Events Using an Extensible Visualization Tool,” presented at the IEEE Signal Processing in Medicine and Biology Symposium, 2017, p. 1.
3. T. van Beelen, “EDFbrowser,” *Teuniz*, 2013. *http://www.teuniz.net/edfbrowser/*.
4. A. Delorme and S. Makeig, “EEGLab,” Swartz Center for Computational Neuroscience, 2002. *https://sccn.ucsd.edu/wiki/EEGLAB*.
5. J. Picone, I. Obeid, and S. Harabagiu, “Automated Cohort Retrieval from EEG Medical Records,” presented at the 26th Conference on Intelligent Systems for Molecular Biology, 2018, pp. 1–7.
6. “Web Playback SDK,” *Spotify for Developers*, 2018. *https://developer.spotify.com/ documentation/web-playback-sdk/*.
7. “HTTP API Guide,” *SoundCloud*, 2018. *https://developers.soundcloud.com/docs/api/guide*.
8. “Hypertext Transfer Protocol – HTTP/1.1,” *Network Working Group*, 2007. *https://www.w3.org/ Protocols/HTTP/1.1/rfc2616bis/draft-lafon-rfc2616bis-03.html*.
9. “Using API Keys,” *Google Inc.*, 2018. *https://cloud.google.com/docs/authentication/api-keys*.