A GNN-based Interactive Application for Earthquake Early Warning

Lindsay Y. Chuang*
EAS, Georgia Tech

Meghana Jain
COC, Georgia Tech

George Jeno
COC,Georgia Tech

Preston T. Lee
COC,Georgia Tech

Hsin-Yi Lin
ISYE,Georgia Tech

Paarth J. Parekh
BIOL,Georgia Tech

1 INTRODUCTION

Earthquake early warning (EEW) systems are crucial applications for high seismic risk areas, such as the U.S. west coast, Japan, Mexico, South Italy, and Taiwan [2, 8, 11, 25, 27]. An EEW system is designed to detect earthquakes and determine their source characteristics (e.g. location and size) by leveraging the very first few seconds of seismic signals (time-series of ground motions). Once an EEW system detects an earthquake, a timely alert can be sent out to users to take necessary actions and reduce damages. The speed and robustness of an EEW system plays a crucial role in minimizing preventable damage, evacuating citizens, and producing a more efficient emergency response [10, 22] - thus, attempts to improve either quality are well-founded in the context of real-world application.

All EEW systems rely on efficient and robust algorithms which automatically process seismic data in real-time. Most of the core algorithms implemented to-date are simple and easy-to-use regression based predictors [18]. Alternative approaches have involved the use of Bayesian probabilistic modeling to predict seismic event parameter probability given input features, allowing for maximum likelihood parameters to be determined [6, 24]. However, there are concerns about the robustness of point estimate extraction from the generated probability distribution, an issue not observed in classification/regression models.

Recent development of deep learning provides alternative ways for fast automation on earthquake related tasks. Convolutional neural networks (CNNs) are one such method, though they can vary in their implementations and outputs. These networks use seismic signals in order to provide source characteristics. Some variations simply output the coordinates of an earthquake's hypo-center [12], and some extend predictions to more detailed characteristics, such as magnitude, location, depth, and origin time [13, 17]. One interesting application comes in the form of DeepQuake [21], a seismo-acoustic event classification CNN. It argues that seismo-acoustic waves can originate not only from earthquakes, but also from events such as volcanic activity, man-made blasts, and the like. With DeepQuake, these differences can be classified, allowing for filtering of earthquake events and others.

Although CNNs are now the most commonly used type of deep learning model in the seismology field, latest researches argue that graph based neural networks (GNNs) are more suitable for earthquake applications [23, 26]. This is because GNNs can better process non-Euclidean structures, which is an inherited data structure of most of the seismic networks worldwide (i.e. seismic networks distributed irregularly spatially). A recent study on GNN demonstrates that GNN can outperform CNN models on automated seismic source characterization task even when only part of the data-set were used [23]. This research opens up a door for potential usage of GNN on EEW system where decisions are made based on incomplete data, though this concept is yet to be proved.

2 OBJECTIVES AND HEILMEIER'S QUESTIONS

This proposal aims to investigate the ability of the GNN model proposed on [23] with a special emphasis of the potential usage for EEW. The GNN model will be probed by various subsets of data with different degree of incompleteness that we create in the future. The predictions made by using these subsets will be carefully compared with the official earthquake catalogs announced by the United States Geological Survey (USGS) in order to measure the accuracy and robustness. The output of the experiments will be organized into a NOSQL database and then being visualized by an interactive web application that's built upon either D3, Leaflet or Bokeh. If successful, this experiment can help with building the next generation of EEW systems. An EEW system with higher efficiency and accuracy can significantly help reduce earthquake hazards in high risk areas. If not successful, the experiment can still shed light on the limitations of GNN. The data structure and the visualizations we will build will still be valuable for other graph based applications.

We plan on first conducting the experiments using 15 years (2000-2015) of earthquake data in Southern California, with an option to extend it to other areas such as the Pacific Northwest. The estimated size of initial data to process is on the order of 500 Gb. This includes ten thousands of earthquakes recorded by 200 seismic stations. We plan on using our local machines to perform the computations, therefore there is practically no costs. The time for conducting the experiments and building the interactive visualizations will both roughly take up to 4-5 weeks, therefore we plan to split the team to two and do both tasks in parallel. More details on how the project will be constructed are provided in section 3. The mid-term and final success will be measured by (1) whether we are able to successfully build the application using the tools listed in section 3, (2) whether the GNN model can be applicable for EEW applications, and (3) Whether the data visualization can clearly showcase the insights of data and easy to use.

3 BUILDING BLOCKS

We plan to use 4 key building blocks to construct this project. Each of the building blocks is described in the following subsections.

3.1 Data scraping

The input data is comprised of two types: (1) seismic waveform data, and (2) earthquake catalog data. Both types of data can be downloaded from the USGS and IRIS datacenter via APIs. We have identified the Obspy package [3] for data scraping and Pandas [15] for initial catalog data storage. List of the tools we plan on using:

- `Obspy data downloader`: A python API used to download seismic data.
- `Pandas`: A python package used to store tabulated data.

3.2 GNN model

We plan on using the GNN model developed in [23] as a bedrock to perform our experiments. The input data will be waveform data, and the output will be predictions on earthquake characteristics. Both the input and output data will be organized into a database (see section 3.3). List of the tools we plan on using:

- `Tensorflow`: A framework for deep learning model training and deployment.
- `GNN model`: A GNN model developed in [23].

3.3 Data structure and database

We plan on using the NOSQL (Not Only SQL) database - MongoDB for data storage. NoSQL databases typically have very flexible schemes. A flexible schema will allow us to easily make changes to the database as the requirements change. Given the high volume and the variety of our data type, the database must be capable enough to cope with the flexibility during operation. MongoDB, as one instance of the NoSQL database, can accommodate a variety of information with different structures into one single geodatabase. Non-relational databases have been widely used in geospatial analysis[4, 20] and we think it's suitable for our application. List of the tools we plan on using:

- `MongoDB`: A source-available cross-platform database program.

3.4 Interactive graph visualization

We have considered several tools that can generate map-like interactive visualizations [16] for our interactive web based application. Among them, Leaflet is a tool with special focus on maps [9]. It has been used in visualizing geographical data [5] and therefore it could help us to build maps with integration of other features. Another tool is D3.js, a powerful and open source library that also supports map visualization[7]. Polymaps is another free tool that uses JavaScript [19] that in our consideration. We will also consider Bokeh, a convenient Python library for building interactive visualizations [1]. Our plan is to evaluate these tools, and then establish the one best suited for our method. List of potential tools:

- `D3.js`: A JavaScript library for visualization.
- `Leaflet, Polymap`: JavaScript libraries with focus on map visualization.
- `Bokeh`: A Python library for interactive visualization.

To help users to easy visualize earthquake detection data, we will provide different views of the data-set. These include but not limited to: (1) density line charts to allow visualization of time-series associated with different stations and earthquakes [14]. (2) graphs showing different relations among data.

4 MEMBER CONTRIBUTION AND PLAN OF ACTIVITIES

All members have participated in two group discussions, helped identified various tools to use, and written fractions of this proposal. In the future, Miss Chuang, Miss Jain, and Mr. Jeno will focus on conducting and designing the experiments. Miss Lin will be in charge of tracking the dataflow and building database. Mr. Lee and Mr. Parekh will focus on designing the interactive visualization.

REFERENCES

[1] Gennady Andrienko, Natalia Andrienko, Steven Drucker, Jean-Daniel Fekete, Danyel Fisher, Stavros Idreos, Tim Kraska, Guoliang Li, Kwan-Liu Ma, Jock Mackinlay, et al. 2020. Big data visualization and analytics: Future research challenges and emerging applications. In *BigVis 2020: Big Data Visual Exploration and Analytics*.

[2] JM Espinosa Aranda, A Jimenez, G Ibarrola, F Alcantar, A Aguilar, M Inostroza, and S Maldonado. 1995. Mexico City seismic alert system. *Seismological Research Letters* 66, 6 (1995), 42–53.

[3] Moritz Beyreuther, Robert Barsch, Lion Krischer, Tobias Megies, Yannik Behr, and Joachim Wassermann. 2010. ObsPy: A Python Toolbox for Seismology. *Seismological Research Letters* 81, 3 (05 2010), 530–533. https://doi.org/10.1785/gssrl.81.3.530 arXiv:https://pubs.geoscienceworld.org/srl/article-pdf/81/3/530/2762059/530.pdf

[4] Huang Q. Cao, K. 2018. Real-time earthquake monitoring with spatio-temporal fields. ISPRS Annals of the Photogrammetry. *Annals of GIS* 24 (2018), 125–138. Issue 22.

[5] A Chourasia, KB Richards-Dinger, JH Dieterich, and Y Cui. 2017. Visual Exploration and Analysis of Time Series Earthquake Data. In *Proceedings of the Practice and Experience in Advanced Research Computing 2017 on Sustainability, Success and Impact*. 1–6.

[6] Georgia Cua, Michael Fischer, Thomas Heaton, and Stefan Wiemer. 2009. Real-time performance of the Virtual Seismologist earthquake early warning algorithm in southern California. *Seismological Research Letters* 80, 5 (2009), 740–747.

[7] Alferd David and Clarence JM Tauro. 2015. Web 3D data visualization of spatio temporal data using data driven document (D3js). *International Journal of Computer Applications* 111, 4 (2015).

[8] Keiji Doi. 2003. Early Warning Systems for Natural Disaster Reduction. Vol. 81. Springer, Chapter 5.6. https://doi.org/10.1007/978-3-642-55903-7_57

[9] Dennis Edler and Mark Vetter. 2019. The simplicity of Modern audiovisual web cartography: an example with the open-source JavaScript Library leaflet. js. *KN-Journal of Cartography and Geographic Information* 69, 1 (2019), 51–62.

[10] Paolo Gasparini, Gaetano Manfredi, and Jochen Zschau. 2011. Earthquake early warning as a tool for improving society's resilience and crisis response. *Soil Dynamics and Earthquake Engineering* 31, 2 (2011), 267–270. https://doi.org/10.1016/j.soildyn.2010.09.004 Prospects and applications of earthquake early warning, real-time risk management, rapid response and loss mitigation.

[11] Richard M. Given, Douglas D.and Allen, Annemarie S. Baltay, Paul Bodin, Elizabeth S. Cochran, Kenneth Creager, Robert M. de Groot, Lind S. Gee, Egill Hauksson, Thomas H. Heaton, Margaret Hellweg, Jessica R. Murray, Valerie I. Thomas, and Douglas Toomey. 2018. Revised technical implementation plan for the ShakeAlert system—An earthquake early warning system for the West Coast of the United States. (2018). https://doi.org/10.3133/ofr20181155

[12] M. Kriegerowski, G. M. Petersen, H. Vasyura-Bathke, and M. Ohrnberger. 2018. A Deep Convolutional Neural Network for Localization of Clustered Earthquakes Based on Multistation Full Waveforms. *Seismological Research Letters* 90 (2018), 510–516. Issue 2A. https://doi.org/10.1785/0220180320

[13] A. Lomax, A. Michelini, and D. Jozinović. 2019. An Investigation of Rapid Earthquake Characterization Using Single-Station Waveforms and a Convolutional Neural Network. *Seismological Research Letters* 90 (2019), 517–529. Issue 2A. https://doi.org/10.1785/0220180311

[14] Dominik Moritz and Danyel Fisher. 2018. Visualizing a million time series with the density line chart. *arXiv preprint arXiv:1808.06019* (2018).

[15] The pandas development team. 2020. pandas-dev/pandas: Pandas. https://doi.org/10.5281/zenodo.3509134

[16] RS Raghav, Sujatha Pothula, T Vengattaraman, and Dhavachelvan Ponnurangam. 2016. A survey of data visualization tools for analyzing large volume of data in big data platform. In *2016 International Conference on Communication and Electronics Systems (ICCES)*. 1–6.

[17] O. M. Saad, A. G. Hafez, and Soliman M. S. 2018. Deep Learning Approach for Earthquake Parameters Classification in Earthquake Early Warning System. *IEEE Geoscience and Remote Sensing Letters* (2018), 1–5. https://doi.org/10.1109/lgrs.2020.2998580

[18] Claudio Satriano, Yih-Min Wu, Aldo Zollo, and Hiroo Kanamori. 2011. Earthquake early warning: Concepts, methods and physical grounds. *Soil Dynamics and Earthquake Engineering* 31, 2 (2011), 106–118.

[19] D SRIDEVI, A KUMARAVEL, and S GUNASEKARAN. 2020. A Review on Big Data Visualization Tools. (2020).

[20] W. Tampubolon. 2014. Utilization of NoSQL database for disaster preparedness. (2014).

[21] L. Trani, G. A. Pagani, J. P. Pereira Zanetti, C. Chapeland, and Evers L. G. 2020. DeepQuake – An application of CNN for seismo-acoustic event classification in The Netherlands. (2020). https://doi.org/10.1002/essoar.10505253.3

[22] PPEW UNISDR. 2006. Global survey of early warning systems: An assessment of capacities, gaps and opportunities toward building a comprehensive global early warning system for all natural hazards. *Platf Promot Early Warn UNISDR—PPEW UN* (2006), 2006.

[23] Martijn PA van den Ende and J-P Ampuero. 2020. Automated Seismic Source Characterization Using Deep Graph Neural Networks. *Geophysical Research Letters* 47, 17 (2020), e2020GL088690.

[24] S Wu, M Yamada, K Tamaribuchi, and JL Beck. 2015. Multi-events earthquake early warning algorithm using a Bayesian approach. *Geophysical Journal International* 200, 2 (2015), 791–808.

[25] Yih-Min Wu, Himanshu Mittal, Ting-Chung Huang, Benjamin M Yang, Jyh-Cherng Jan, and Sean Kuanhsiang Chen. 2019. Performance of a low-cost earthquake early warning system (P-Alert) and shake map production during the 2018 Mw 6.4 Hualien, Taiwan, earthquake. *Seismological Research Letters* 90, 1 (2019), 19–29.

[26] Keisuke Yano, Takahiro Shiina, Sumito Kurata, Aitaro Kato, Fumiyasu Komaki, Shin'ichi Sakai, and Naoshi Hirata. 2020. Graph-partitioning based convolutional neural network

for earthquake detection using a seismic array. (2020).
[27] Aldo Zollo, Giovanni Iannaccone, Vincenzo Convertito, Luca Elia, Iunio Iervolino, Maria Lancieri, Anthony Lomax, Claudio Martino, Claudio Satriano, Emanuel Weber, and Paolo Gasparini. 2011. *Earthquake Early Warning System in Southern Italy.* Springer New York, New York, NY, 175–201. https://doi.org/10.1007/978-1-4419-7695-6_13