Community Sentiment Analysis for Identifying Social Vulnerability Following Earthquakes

Principal Investigator
Sifat Muin, University of Southern California

Research Team
- Kristina Lerman, Information Science Institute, University of Southern California

Start-End Dates:
08/01/2023 - 07/31/2024

Abstract

Earthquake resilience, a critical facet of disaster preparedness, traditionally focuses on structural and infrastructural integrity, often overlooking the human dimension of recovery. This study aims to bridge this gap by proposing a novel framework that harnesses the power of social media and natural language processing (NLP) techniques for sentiment analysis to gauge communities' emotional responses and attitudes before, during, and after an earthquake event. The primary objective is to comprehend the diverse responses of communities to seismic events, and identify vulnerabilities.

This research consists of two key components. *Sentiment Expression Analysis* employs advanced natural language processing techniques, prominently featuring the Robustly Optimized BERT Approach (RoBERTa) Model for Sentiment Analysis, to investigate sentiments and emotions within post-earthquake community responses. *Correlation Exploration* aims to uncover complex links between social vulnerability and certain population traits. This is achieved by merging insights extracted from sentiment analysis with demographic statistics obtained from census data. This methodology allows us to gain a deeper understanding of these connections and their impact on various groups of people.

Employing the 2019 Ridgecrest earthquake sequence as a case study, the proposed framework will conduct a thorough community sentiment analysis, thereby shedding light on the emotional landscape of affected communities. The integration of census and reconnaissance data will further enhance the depth of this analysis. Additionally, topic analysis will be employed to discern specific issues and concerns voiced by communities in need.

By merging perspectives from earthquake engineering, data science, and demographics, this research aspires to construct a holistic framework dedicated to augmenting earthquake resilience within our communities. This endeavor seeks to contribute to the comprehensive understanding of earthquake preparedness and response, with the ultimate goal of fostering more resilient communities that are better equipped to withstand the impact of earthquakes.

Deliverables

The main deliverable of this project will be a PEER report. Publication of several posters, conference presentations, and journal papers that provide a comprehensive overview of the research progress and noteworthy findings is anticipated. Specifically, maps will be provided to highlight the correlation between the sentiment scores and the various social and hazard threats that identify vulnerable communities within a given California county. Moreover, a GitHub repository will be created, and all collected data and trained models will be available there.

Research Impact

This research project is set to rigorously investigate the intricate relationship between human sentiment, earthquake intensity, damage, and the social environment. It is expected to yield significant outcomes, including:

1. Precisely identifying communities with heightened vulnerability, allowing for precise resource allocation and targeted support.
2. Substantially advancing disaster preparedness by providing deep insights into how communities respond to seismic events, thereby enabling the enhancement of preparedness and response strategies.
3. Serving as a foundation for evidence-based policymaking, facilitating the development of comprehensive disaster resilience measures.
4. Pioneering technological advancements through the application of state-of-the-art analytical tools, with the potential to benefit various domains beyond earthquake engineering
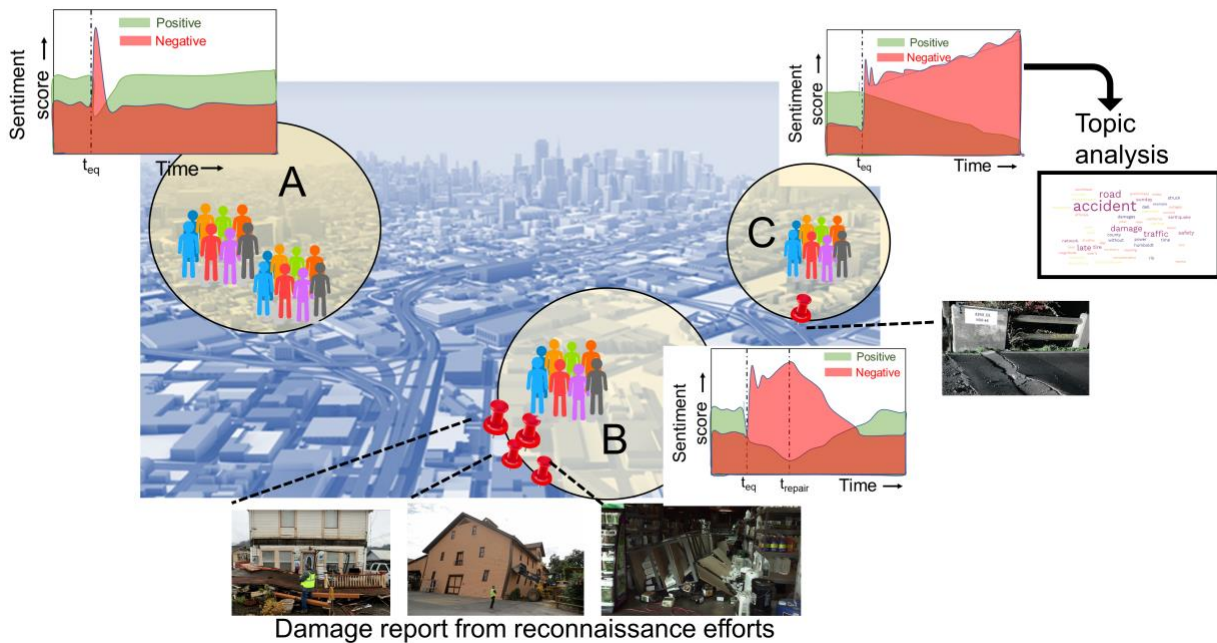
Project Image

Figure 1 Impact of an earthquake on different communities which is tracked by sentiment analysis enabling identification of important issues and disparities. Community A had minimal impact. Community B shows growing negative sentiments due to significant damage, but once retrofit efforts are taken there, it satisfies the community. Community C experienced moderate damage to an important road that has not been addressed, and consequently, negative sentiment increased over time.
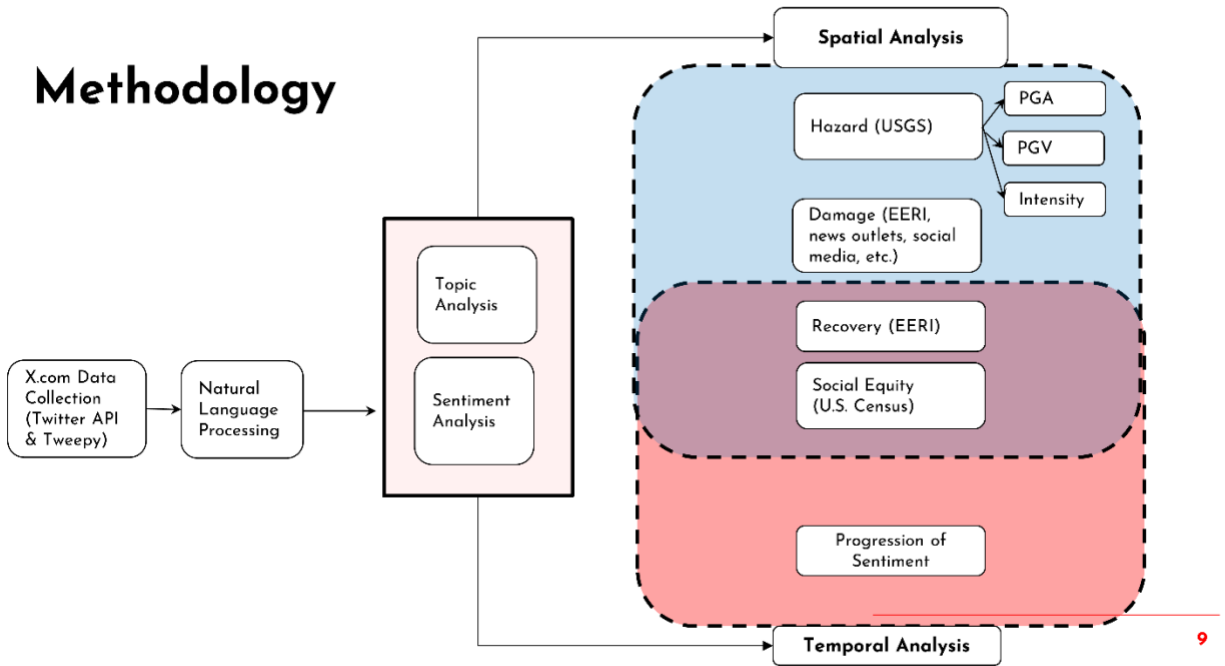
If needed, additional Project Image

Figure 2: Proposed methodology of the project