Portfolio Project: Dynamic Risk Visualization in Abila

Antarang Poogalia
School of Computing and Augmented Intelligence, ASU
Tempe,USA
apoogali@asu.edu

I. INTRODUCTION

The VAST Challenge 2021: Mini-Challenge 3, posed a unique and complex problem centered around dynamic risk visualization in the city of Abila. The scenario unfolded on January 23, 2014, and required participants to retrospectively analyze and identify risks using limited data. Our team, comprising Vibhor Agarwal, Rahul Ashwin Sheth, Vaibhav Somani, Antarang Poogalia, Manohar Veeravalli, and Laksh Gangwani, aimed to develop a solution that could not only identify risks but also propose effective mitigation strategies. The primary challenge was to analyze a singular data stream composed of microblog records and emergency dispatch texts from local police and fire departments. Additionally, we utilized maps of Abila and other relevant background documents, aiming to understand and evaluate changing risk levels to the public over time.

II. EXPLANATION TO THE SOLUTION

Our solution was a series of integrated visual analytics tools designed to handle and interpret the complexity of the data effectively. We developed a range of visualizations, each serving a specific purpose in our analysis:

1. Task 1 (Characterizing Content Types):

The goal was to differentiate meaningful event reports from general chatter, junk, or spam in the dataset.

Stream Graph: This visualization was employed to track the volume and flow of conversations over time, which helped in identifying significant spikes that often correlated with important events. This tool was key in distinguishing meaningful content by showing when and how conversation patterns changed.

Circle Packing: This chart was used to analyze the distribution of messages from individual users. It helped in identifying those users who were contributing substantial information, as opposed to those generating irrelevant content or spam. The volume of messages from each user was visually represented, facilitating the identification of key informants and the exclusion of less relevant or spammy data.

2. Task 2 (Evaluating Risk Evolution):

The aim was to visually represent and assess how the risk to the public changed over the course of the evening, considering potential consequences and the number of people affected.

Network Graph (Link Nodes): This graph demonstrated the interconnections and context among various messages, categorizing them into distinct risk-related themes. It was critical in understanding the clustering of messages and their risk implications.

Beeswarm Chart: Illustrated the density of messages over time, highlighting periods with high activity which could indicate increased risk.

Word Cloud: Provided a visualization of the most frequent words in messages during selected time frames, giving insights into prevailing themes and concerns, which were crucial for risk assessment.

3. Task 3 (First Responder Deployment Decision-making):

This task involved determining the optimal location to send a team of first responders based on the analysis.

Dynamic Map with Pulsating Effects: The map displayed incidents with geographical coordinates and assigned risk scores. The pulsating effect on the map indicated the intensity of risk, guiding decision-making for first responder deployment. It was essential for identifying the most critical incident at any given time. The map also classified events into categories (e.g., police, vehicle accident, fire) and included a time control feature for real-time analysis.

III. RESULTS

The results of our visualization tools provided a comprehensive and multi-faceted analysis of the unfolding situation in Abila. Each visualization tool contributed uniquely to our overall understanding:

1. **Stream Graph:** The Stream Graph was instrumental in visualizing the flow and intensity of communications over time. It revealed patterns in the data that were not immediately apparent, such as spikes in communication during key events. This visualization was particularly effective in filtering out irrelevant data, allowing us to focus on significant temporal patterns and correlations between the occurrence of events and public reactions.

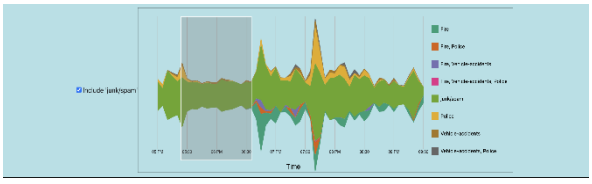

Fig. 1 : Stream Graph

2. **Circle Packing Chart:** The Circle Packing Chart offered a novel perspective on the distribution and frequency of messages by individual authors. This visualization was key in identifying influential communicators and information spreaders within the data. It enabled us to understand who the key players were during the crisis, based on the volume and impact of their messages.

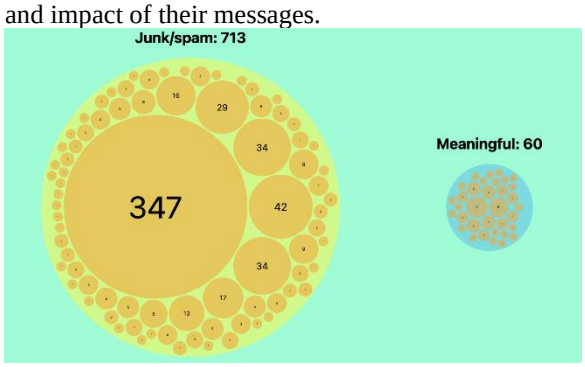

Fig. 2 : Circle Packing Chart

3. **Network Chart:** The Network Chart was a powerful tool for uncovering the hidden relationships between different pieces of information. By analyzing the connections and associations between various nodes, we were able to identify potential leads and insights that would have been challenging to discern otherwise. The risk scores attached to each node provided a quick and intuitive way to gauge the significance and impact of each message.

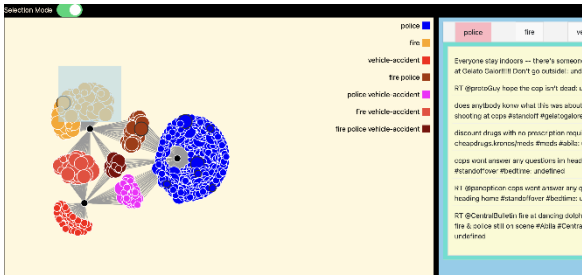

Fig. 3 : Network Chart

4. **Beeswarm Chart:** The Beeswarm Chart offered a dynamic view of the chronological progression of events, highlighting the density and intensity of activities at particular moments. This was crucial in understanding the pace at which situations escalated and the corresponding public response. The ability to focus on specific time frames allowed for a detailed analysis of critical moments.

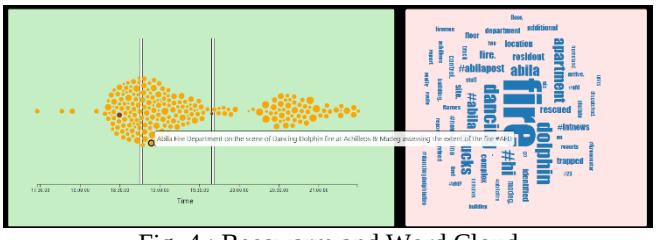

Fig. 4 : Beeswarm and Word Cloud

5. Word Cloud Observations: The Word Cloud gave us qualitative insights into communication during the crisis. By analyzing the most frequently used words in a selected timeframe, we could identify the main themes and concerns expressed by the public. This visualization was particularly useful in understanding public sentiment and the most pressing issues during specific periods of the crisis.

6. Geospatial Visualization on Abila's Map: Our geospatial visualization provided a real-time depiction of incidents across the city of Abila. The use of distinctive icons for different types of incidents (accidents, emergencies, fire events) and their pulsating effects based on risk levels offered an immediate visual assessment of the severity and distribution of incidents. This aspect of our solution was pivotal in proposing strategic locations for deploying emergency response teams.

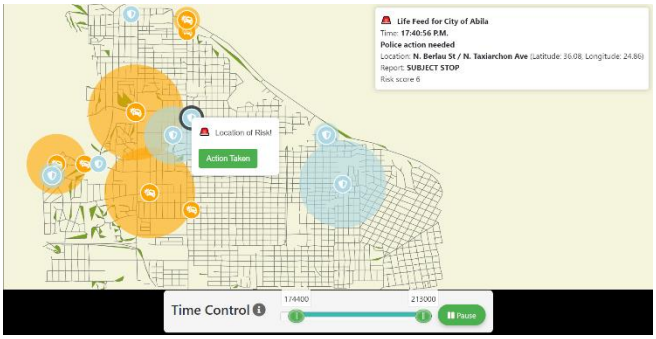

Fig 5 : Geospatial Map Visualization with pulsating effects

Overall, the combined use of these visual analytics tools allowed us to not only track and analyze the events as they unfolded but also to provide actionable insights for better crisis management. Our analysis revealed key patterns, high-risk periods, and critical communication nodes, which could be invaluable for emergency response teams and city planners in managing similar situations in the future.

This comprehensive and detailed understanding of the events in Abila showcases the potential of visual analytics in managing and responding to urban crises, marking a significant step forward in the field of data-driven emergency management.

IV. CONTRIBUTIONS AND LESSONS LEARNED

As part of the group project, my contributions were multifaceted and integral to its success:

1. Data Preprocessing and Classification:
   - I spearheaded the preprocessing of the dataset, which included classifying untagged messages. Using algorithms such as TF-IDF, I categorized the messages, laying the groundwork for further analysis.
   - I conducted an in-depth study of the dataset to comprehend the underlying problem and articulated the tasks necessary for addressing the project's objectives. This process

involved defining labels for the data, ensuring that these labels were consistently applied across all datasets and used effectively in all three tasks.

2. Visualization Ideas:
- I was instrumental in suggesting innovative visualization approaches. The idea of using circle packing to represent the author and message count, as well as the concept of a dynamic map with a pulsating effect, were initiatives I proposed. These ideas were crucial in enhancing the project's data representation capabilities.

3. Map of Abila and Risk Assessment:
- I meticulously reviewed the emergency call center data to ensure the accuracy of risk scores, a critical aspect for the real-time update feature of the map.
- To compensate for the absence of geographical coordinates in the dataset, I developed a technique to extract latitude and longitude from the provided geojson data based on street locations, assigning these coordinates to each relevant data entry.
- I advocated for highlighting only high-risk incidents on the map. By aggregating risk scores and adjusting the visual representation (size and pulsation of icons) for events within close geographic and temporal proximity, I ensured that the map remained informative without being cluttered.

Acquired Skills and Knowledge

This project was a substantial learning experience, contributing to my professional growth in several key areas:

1. Front-End Application Development with Data:
- I learned how to develop a front-end application that seamlessly integrates with large datasets, enhancing my skills in interactive web development.

2. Handling Special Cases in Data:
- The project taught me strategies for dealing with missing or misaligned data, a common challenge in real-world data analysis.

3. Leadership and Task Distribution:
- I honed my leadership skills by effectively distributing tasks among team members, ensuring efficient progress and cohesive teamwork.

4. Data Preprocessing Importance:
- I gained a deeper understanding of the importance of data preprocessing. This process is crucial for ensuring data quality and reliability, which are fundamental for accurate analysis and visualization.

5. Proficiency in D3.js Library:
- My experience with this project expanded my knowledge and skills in using the D3.js library, a powerful tool for creating sophisticated data visualizations.

These skills and experiences not only contributed to the success of this project but also enhanced my overall expertise in data visualization and analysis.

V. References

[1]https://marketplace.visualstudio.com/items?itemName=ritwickdey.LiveServer.
[2]https://observablehq.com/@d3/streamgraph.
[3]https://observablehq.com/@d3/force-directed-graph-component.
[4]https://d3-graph-gallery.com/wordcloud.html.
[5]https://www.d3indepth.com/geographic/putting-it-all-together.
[6]https://marketplace.visualstudio.com/items?itemName=ritwickdey.LiveServer.