Saving St. Himark – Vast Challenge 2019

TNM098 – Advanced Visual Data Analysis

Algot Sandahl David Robín Karlsson

June 4, 2021

1 Task and data

This project report details visualisations and analyses used to solve the *Vast Challenge 2019 – Mini-Challenge 1* [1]. The goal is to utilise user-reported damage of earthquake-activity to guide the city's emergency response.

1.1 Data

The dataset contains user-reported data regarding earthquake shake intensity, infrastructural damage, the location in the city the report regards, and the date and time of when the report was created. The infrastructural damage is split up into different categories:

- Sewer and water
- Power
- Roads and bridges
- Medical
- Buildings

All reported data use a discrete 0 to 10 scale, where 0 is no damage and 10 is extensive damage. As such, a single row of data in the dataset takes the form shown in table 2, where the integer under location indicates which borough the report regards. The data also contains rows where some ratings are missing.

The report distribution of each area, from a total of 83070 reports, is shown in table 1.1.

The data also includes a brochure of sorts, detailing city layout, different facts and benefits of moving to St. Himark which can be used in solving the task. For example, it contains information regarding the main types of estate in the different locations. It also contains a map of the different boroughs with additional information. This map is shown in figure 1.

Table 1: *Number of reports from each location in St. Himark.*

Location	Reports	% of all reports
Scenic Vista	13889	16.7
Old Town	13535	16.3
Broadview	8796	10.6
Easton	6429	7.7
Northwest	5390	6.5
Weston	4729	5.7
East Parton	4545	5.5
Pepper Mill	3463	4.2
Downtown	3160	3.8
Southton	3145	3.8
Safe Town	2919	3.5
Terrapin Springs	2844	3.4
Chapparal	2213	2.7
Oak Willow	1727	2.1
Southwest	1710	2.1
Palace Hills	1662	2.0
West Parton	1515	1.8
Cheddarford	1226	1.5
Wilson Forest	173	0.2

2 Analysis strategies

This section explains the methods used to allocate the rescue resources

2.1 Population distribution

Since the given data provides no information to the distribution of the population through the different boroughs, the total amount of reports from an area can be used as attain an approximate population distribution. The distribution of reports is shown in table 1.1 and visualised in figure 2 which shows that approximately half the population lives in the four most populated boroughs.

The population distribution is important in deciding where to allocate resources to maximise the number of people getting help.

time	sewer_and_water	power	roads_and_bridges	medical	buildings	shake_intensity	location
2020-04-08 17:50	10.0	6.0	10.0	3.0	8.0	0	1

Table 2: *Example of a row in the data.*

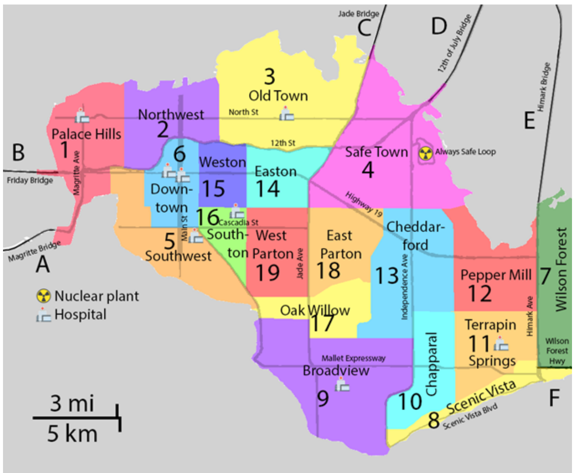

Figure 1: *The map given with the data [1].*

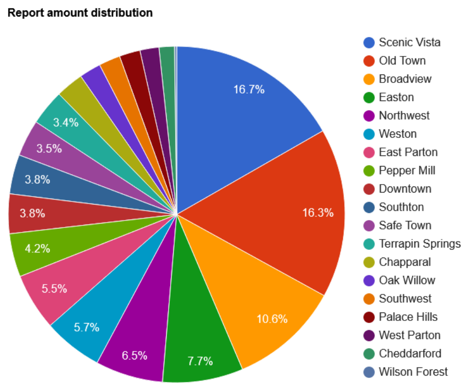

Figure 2: *The distribution of the number of reports.*

2.2 Infrastructure

The city brochure mentions important and earthquake-sensitive infrastructure located in different boroughs. Perhaps the most important is the Always Safe Nuclear Power Plant in Safe Town. It provides 72% of the city's power, power that will be paramount in the emergency response. There can be no skimping on resources to ensure the power plant's continued operation and prevention of a nuclear disaster. Therefore, Safe Town is a high priority borough.

Arguably the second most important infrastructure in an earthquake event are the hospitals. The city has eight operating hospitals specialising in different areas. Ideally all hospitals would be kept fully operational during an earthquake-event but if the resources do not allow it some hospitals will be prioritised. Downtown houses the Trauma Hospital, specialised in trauma- and burn care which are key medical services in an earthquake-event. Downtown also houses the Children Hospital which must be prioritised as children will inadvertently be affected by an earthquake. Downtown's neighbouring borough Southton houses the Eaglepeak Hospital specialised in critical care, another important medical service in an earthquake-event. The neighbouring borough Southwest houses an unnamed hospital which can be used to offload the previously mentioned hospitals. This forms a hospital core of Trauma Hospital, Children Hospital, Eaglepeak Hospital and the unnamed hospital in Southwest to which patients will be ferried in the event that not all hospitals can stay fully operational. This means that the boroughs Downtown, Southton and Southwest are prioritised.

2.3 Visual analytics

While the previous two sections presents data that is generally useful during the whole duration of the earthquake event is it difficult to discern how the situation evolves over time. Therefore, some visual analytics of the data is utilised to be enable snap decisions and counter situations that arise locally in time.

These visualisations were implemented as a dashboard using d3.js. The reported damages from all boroughs were visualised in full using heatmap matrices, as seen in figure 4. The borough and the number of reports from it is displayed next to its heatmap matrix.

To allow an overview of all reports, the option to plot the average value of all reported damage at a single time point is given (severity), as seen in figure 5.

It is possible to plot only a certain parameter from the reports to analyse how for example the damage to roads and bridges evolve. This removes visual clutter from the full report while giving more detailed detailed information than the average plot.

These visualisations allow analysis of how the consequences of the earthquake-event evolves over time and details what infrastructure has been affected in at certain periods. This is a great tool in protecting important infrastructure in the prioritised boroughs

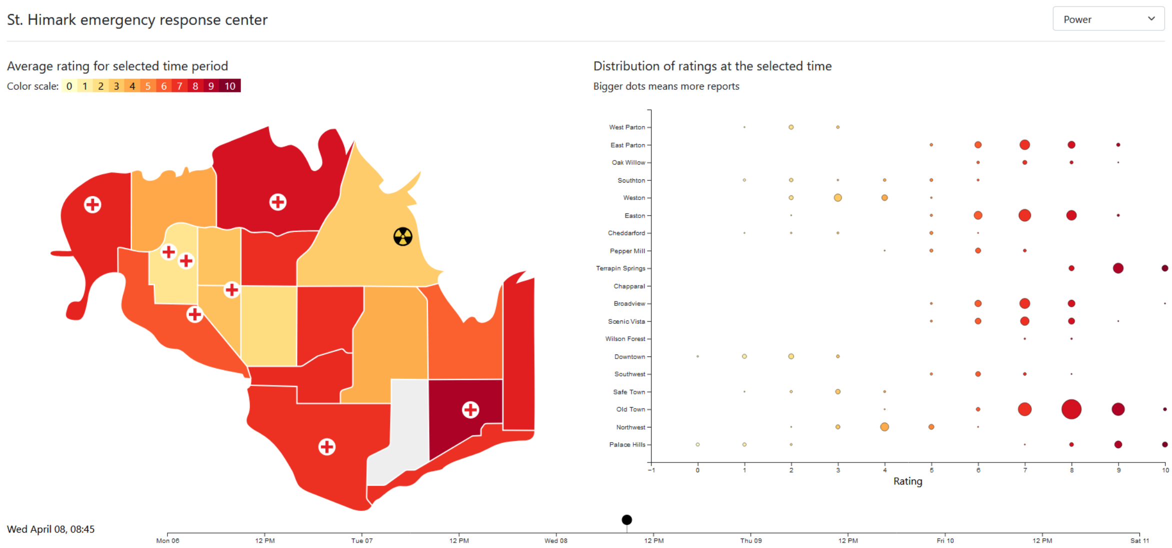

Figure 3: *The cloropleth map and bubble chart showing the damage to power infrastructure.*

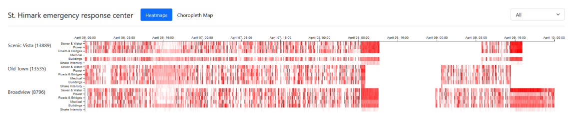

Figure 4: *The full reports of three different locations.*

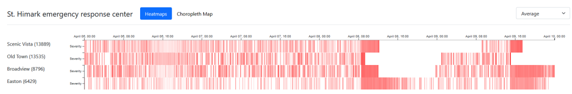

Figure 5: *The averaged reports of three different locations.*

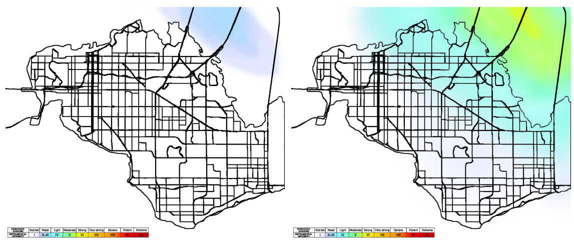

Figure 6: *Shakemap of the pre-quake (left) and large quake (right). Part of the data [1].*

presented above.

The heatmap matrices present a weakness in poor temporal granularity. The time-axis ticks increases in steps of eight hours which does not allow for minute-to-minute decisions which might be crucial in an emergency. Therefore, a second visualisation in form of a cloropleth map showing the average severity of a borough at a certain time was implemented. The map has an accompanying time-slider, moving in ticks of five minutes, allowing selection of a specific date and time. This visualisation possesses temporal granularity which matches the given data (five minute intervals) and allows for momentaneous analysis of the current situation which supplies the data needed for informed snap decisions.

Accompanying the choropleth map is a bubble chart which plots the number of reports with each rating for each of the locations. By visualising the number of reports from an area, the uncertainty of the data can be estimated, i.e if there are very few reports detailing extensive damage in a location the data is not as reliable as if it had been multiple reports. Additionally it can be seen how the reports are distributed and outliers can be identified. The choropleth map and bubble chart visualise reported damage from a single variable in the data, selectable with a drop-down menu, and can be seen in figure 3.

2.4 Shakemaps

The dataset also provided two shakemaps, seen in figure 6, detailing a small pre-quake and a significantly larger earthquake later on.

With all the presented data and visual analytic tools, the authors are ready to guide an informed and optimal emergency response.

3 Resource allocation

With the prioritised areas, population distribution and visual analytics presented above, an informed

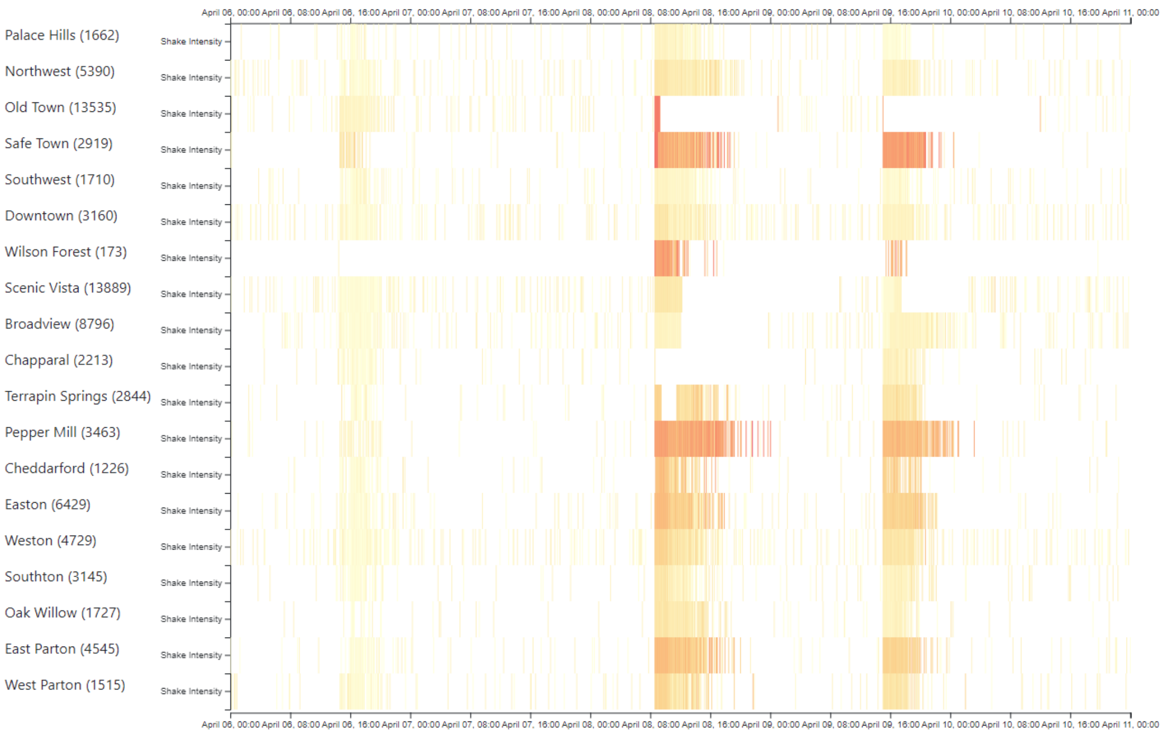

Figure 7: *The reported shake intensity over time.*

plan on how to spend emergency resources can be made. As mentioned in the brochure, the city has 34 different fire stations, with no explicitly mentioned locations, capable of dispatching emergency aid.

The data duration stretches five days, April 6 00:00 to April 10 00:00. However, figure 7 clearly shows that there are three quakes: a foreshock followed by two major shakes.

3.1 The foreshock

As seen in figure 6, the epicentre of both quakes is off the northeastern coast. In the initial pre-quake small parts on the coasts of Old Town and Safe Town are affected with a weak to light intensity. Using the cloropleth map and bubble chart we can that this event occurs around 16:00 on April 6 where a small spike in reports coming from Old Town can be seen. The spike of reports and reported intensity can be seen in the heatmaps as well. These two effects can be seen in figure 9. There are also a non-trivial amount of reports coming from Downtown, Scenic Vista, Broadview, Chapparal and Weston regarding very light damage to power, sewer and water.

Since Old Town is the second most populated borough, the nuclear power plant is located close-by in Old Town and the other boroughs are reporting very light to no activity or damage to infrastructure, the emergency resources should mostly be allocated to Old Town and Safe Town while smaller teams spread around the city to investigate the very small reported damage to infrastructure.

3.2 The first major shock

At around 9:00 on April 8 the first widespread shake in the city is recorded. Old town reports severe shake intensities while many others report slightly lower but still severe intensities, which can be seen in figure 10. Severe damage to all infrastructure is reported in the affected boroughs with power being hit the hardest, which can be seen in figure 8.

Resources need to be focused mainly to Old Town, Safe Town, Scenic Vista, Broadview and Easton with a smaller team to the much less populated Terrapin Springs. Chapparal most likely need attention as well although there is not much data to base this decision on because of a power outage. Other small teams can be sent to less affected boroughs and survey teams should inspect boroughs with no reports in case power outages affect reporting, in which case resources can be shifted if they are needed elsewhere.

Figure 8: *Heatmaps of all reported variable data from half of all areas.*

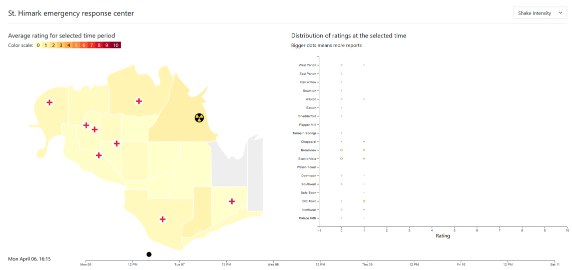

Figure 9: *Reported data spikes from the pre-quake occuring on April 6, around 16.00.*

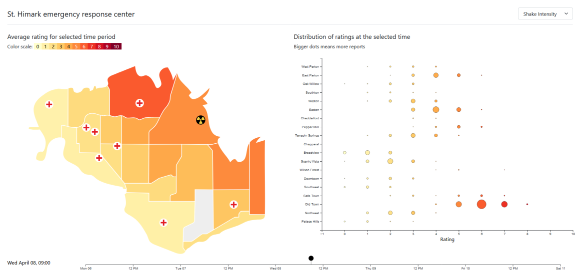

Figure 10: *Reported shake intensity data from the first major shake on April 8, around 9:00.*

3.3 The second major shock

The second major shock occurred around 15, April 9. A power outage hit Old Town almost instantly after this shake, but a substantial amount of reports of high damage were received the minutes before the outage; the roads and bridges rating is shown in figure 11. From the heatmap of all variables, it can also be seen that Scienic Vista has substantial damage in most categories in spite of a moderate shake intensity.

Furthermore, Broadview and Chapparal has high ratings in multiple categories. Southton, Easton, Terrapin Springs and Safe town has damage in many categories but primarily in one, and should be prioritised after the locations with more damage.

4 Discussion and future work

One of the major drawbacks of the dashboard is that the uncertainty of the data can not be assessed from the heatmaps. The information can be found using the bubble chart visualisation, but that can only be done for small time intervals, and the changes in time can be difficult to spot. This problem is further amplified by the vast changes in report density throughout the duration of the event.

One possible solution to the problem of the varying report density is to allow for varying time step lengths that ensure there is enough information to make calculations from, in each interval.

To solve the problem of the uncertainty changes over time, it might work to fit a model to the data and

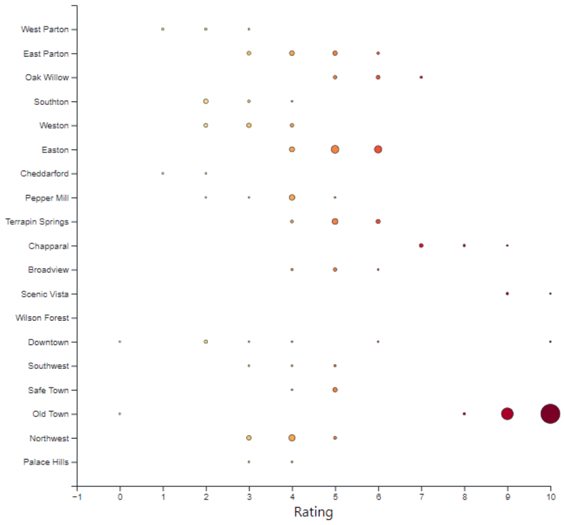

Figure 11: *Reported ratings for road and bridges for April 9, 15:00*

to calculate an uncertainty measurement that can be visualised together with the data in the heatmaps, for example using a two-dimensional colour scale.

References

[1] VAST challenge. (2019). "Mini-challenge 1: Crowdsourcing for situational awareness." Accessed 2021-05-17, [Online]. Available: https://vast-challenge.github.io/2019/MC1.html.