VIS 2019

Interactive Ranking Uncertain Multivariate Ordinal Time Series: Citizen Science and Uncertainty

MC1 Award for Excellent Quantification of Abnormalities

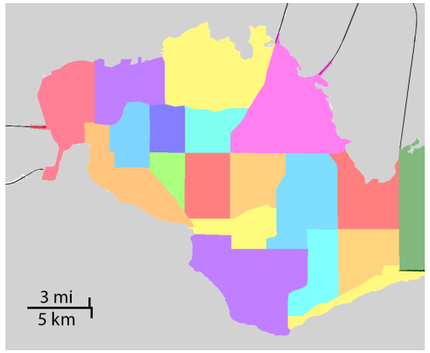
Shichao Jia, Jiaqi Wang, Zeyu Li, and Jiawan Zhang
Tianjin University

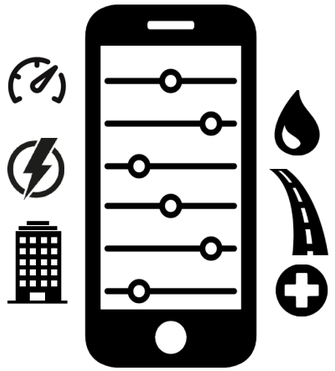
Agenda

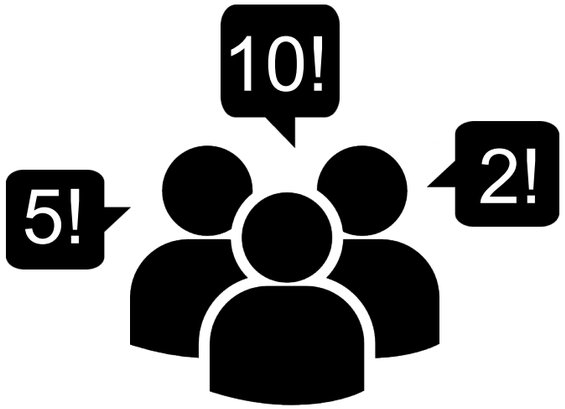
- Data
- Questions & Task Abstraction
- Visual Design
- System
- Insights
- Conclusion

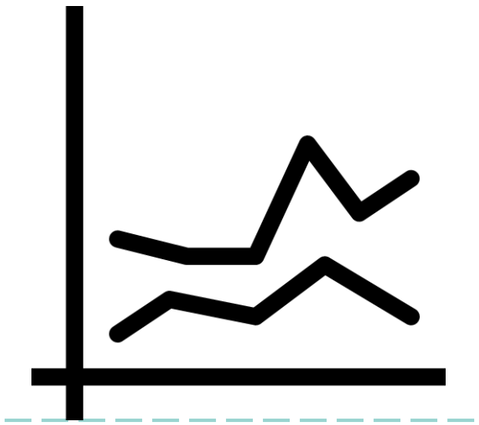
Data

- 19 locations → multivariate (geographic)
- 6 dimensions → multivariate (descriptive)
- 11 levels (0 - 10) → ordinal
- Survey data → uncertain
- Every 5 minutes → time series
- **Summary**: Uncertain Multivariate Ordinal Time Series



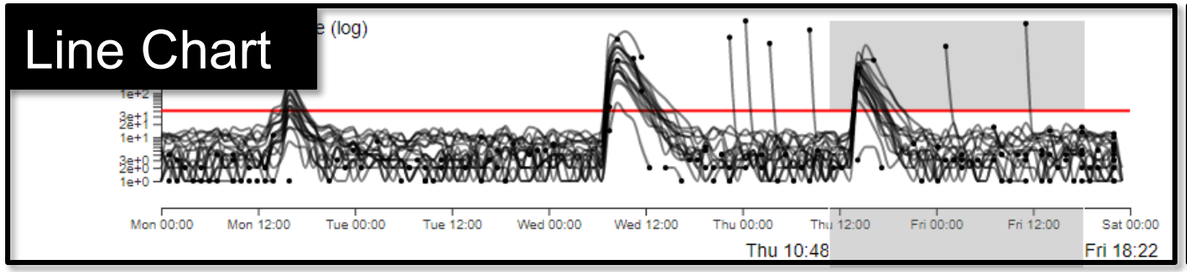


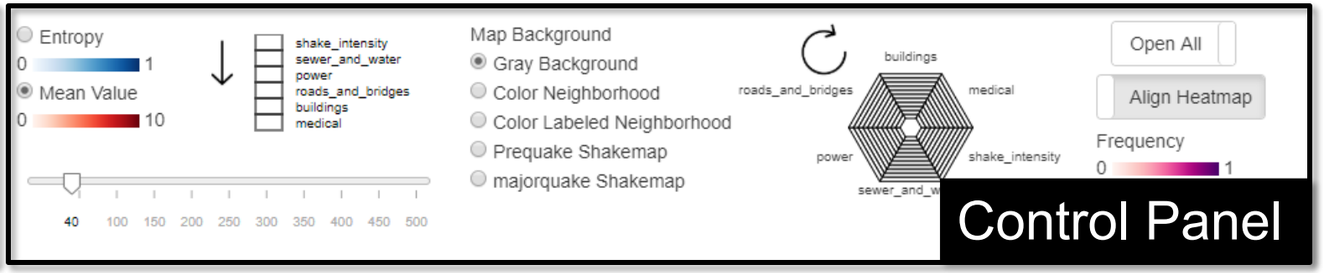


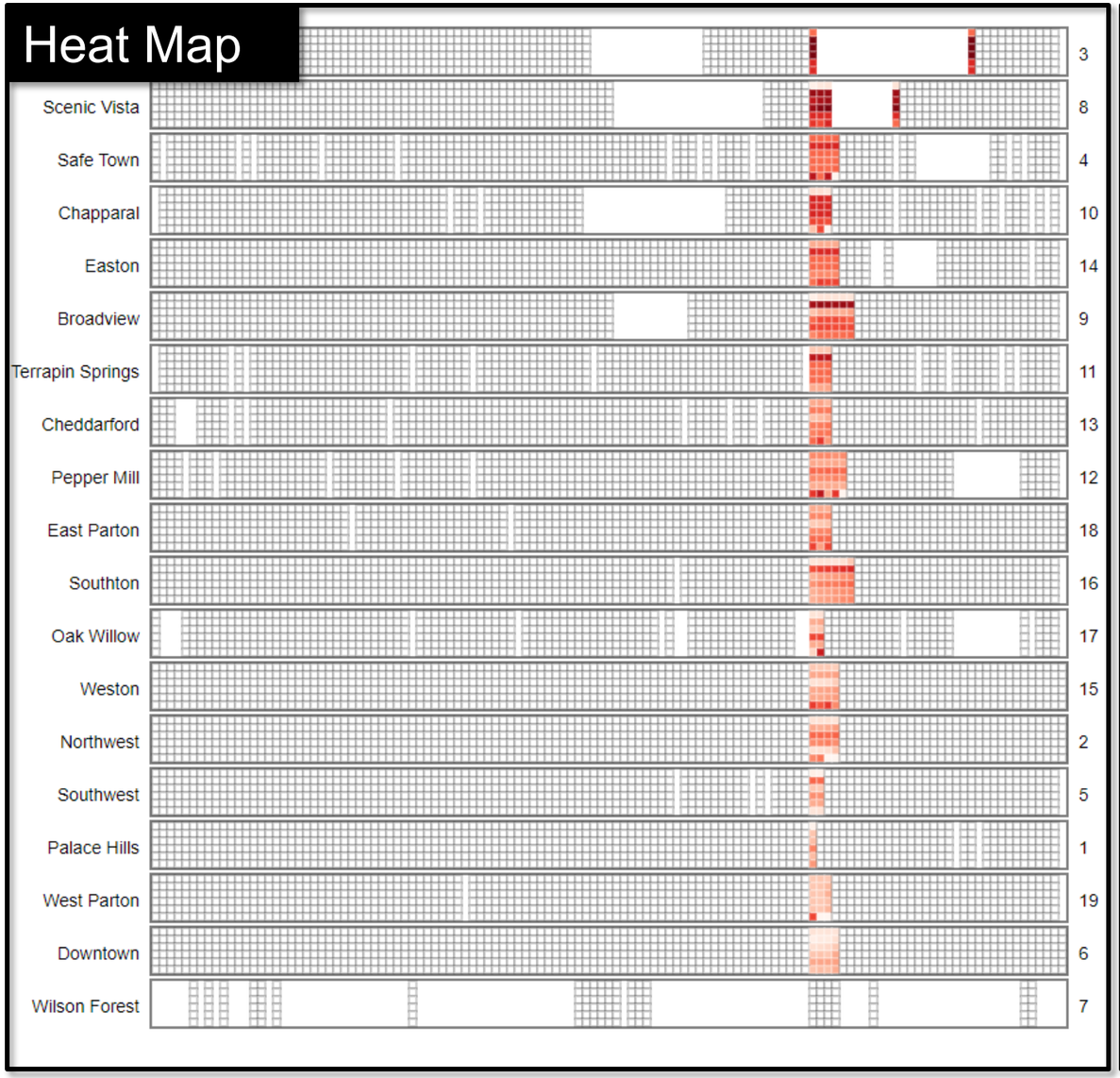


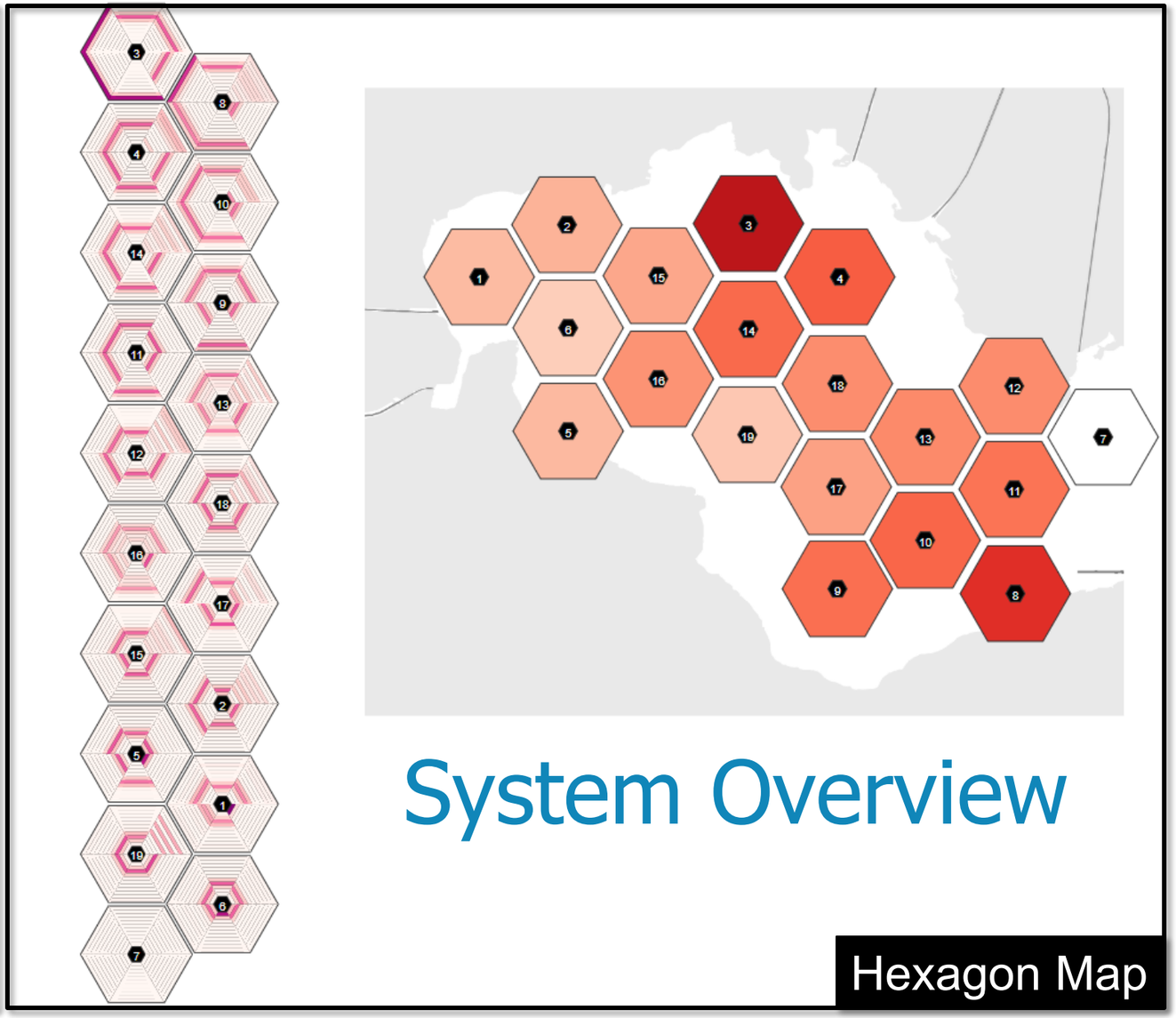
Questions & Task Abstraction

- Task for Q1: Prioritize neighborhoods based on different dimensions
- Task for Q2: Prioritize neighborhoods based on uncertainty
- Task for Q3: Visualize uncertain multivariate ordinal time series


Line Chart
(log)
1e+2
3e+1
1e+1
3e+0
2e+0
1e+0
Mon 00:00
Mon 12:00
Tue 00:00
Tue 12:00
Wed 00:00
Wed 12:00
Thu 00:00
Thu 12:00
Fri 00:00
Fri 12:00
Sat 00:00
Thu 10:48
Fri 18:22

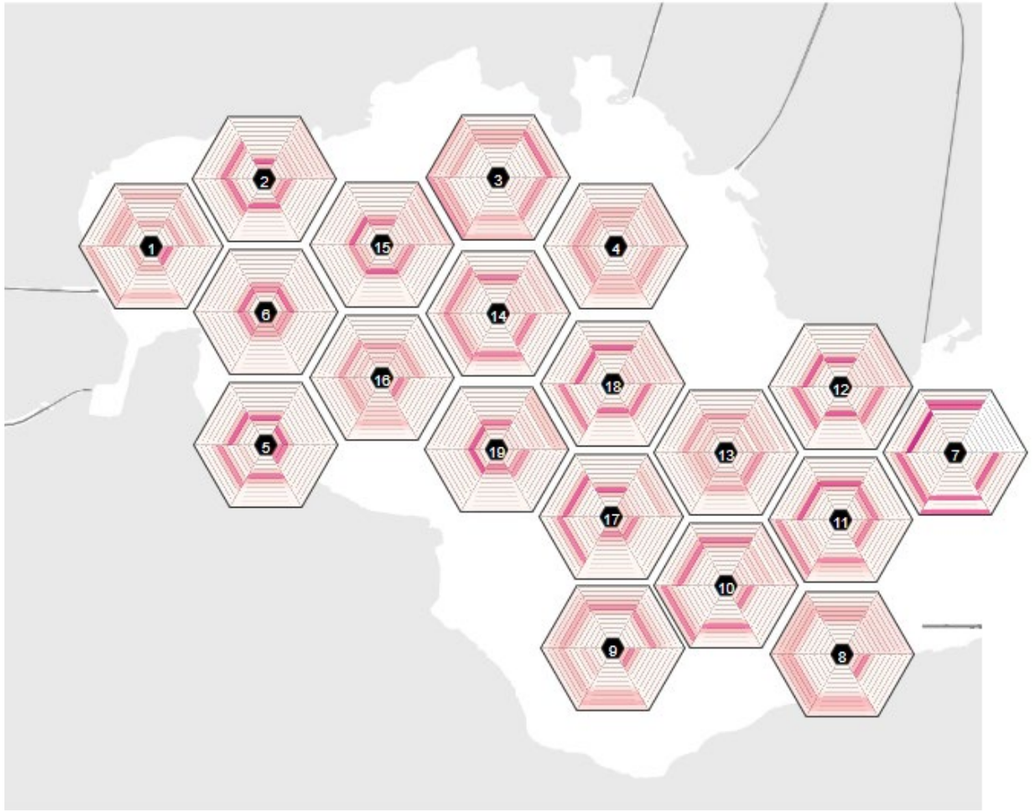

Entropy
0
1
Mean Value
0
10
shake_intensity
sewer_and_water
power
roads_and_bridges
buildings
medical
40
100
150
200
250
300
350
400
450
500
Map Background
Gray Background
Color Neighborhood
Color Labeled Neighborhood
Prequake Shakemap
majorquake Shakemap
buildings
roads_and_bridges
medical
power
shake_intensity
Open All
Align Heatmap
Frequency
0
1
Control Panel

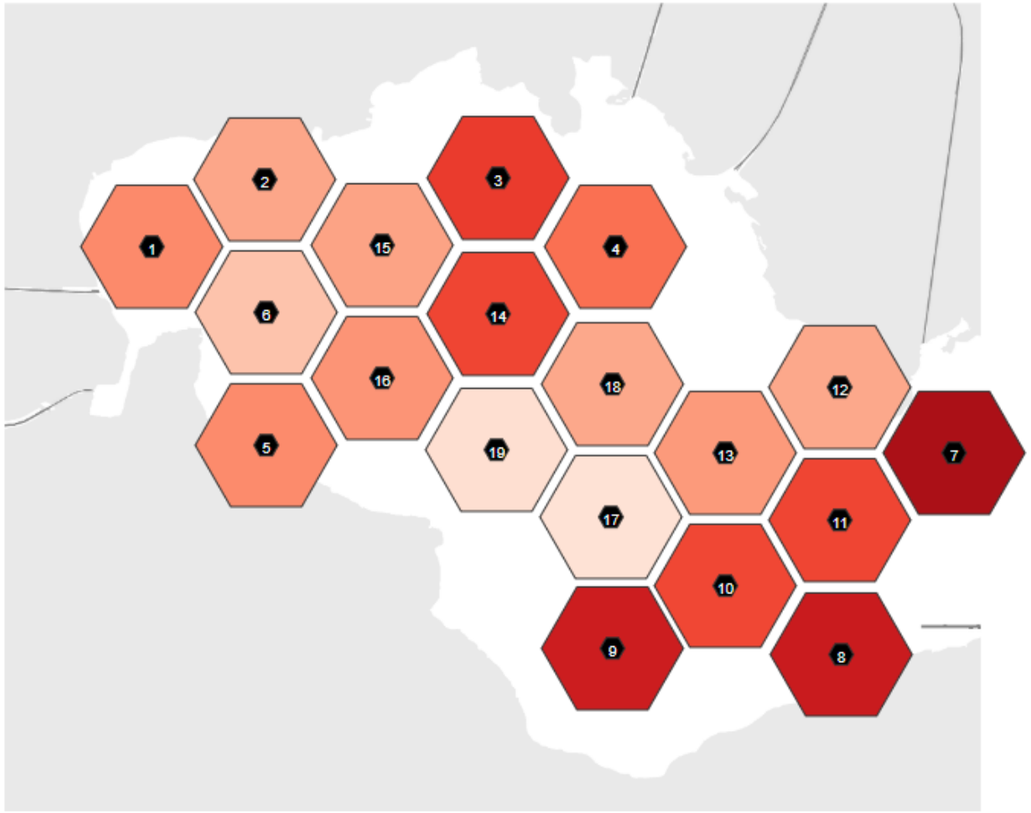

Heat Map
Scenic Vista 8
Safe Town 4
Chapparal 10
Easton 14
Broadview 9
Terrapin Springs 11
Cheddarford 13
Pepper Mill 12
East Parton 18
Southton 16
Oak Willow 17
Weston 15
Northwest 2
Southwest 5
Palace Hills 1
West Parton 19
Downtown 6
Wilson Forest 7
3


System Overview
Hexagon Map

Visual Design: Hexagon Map

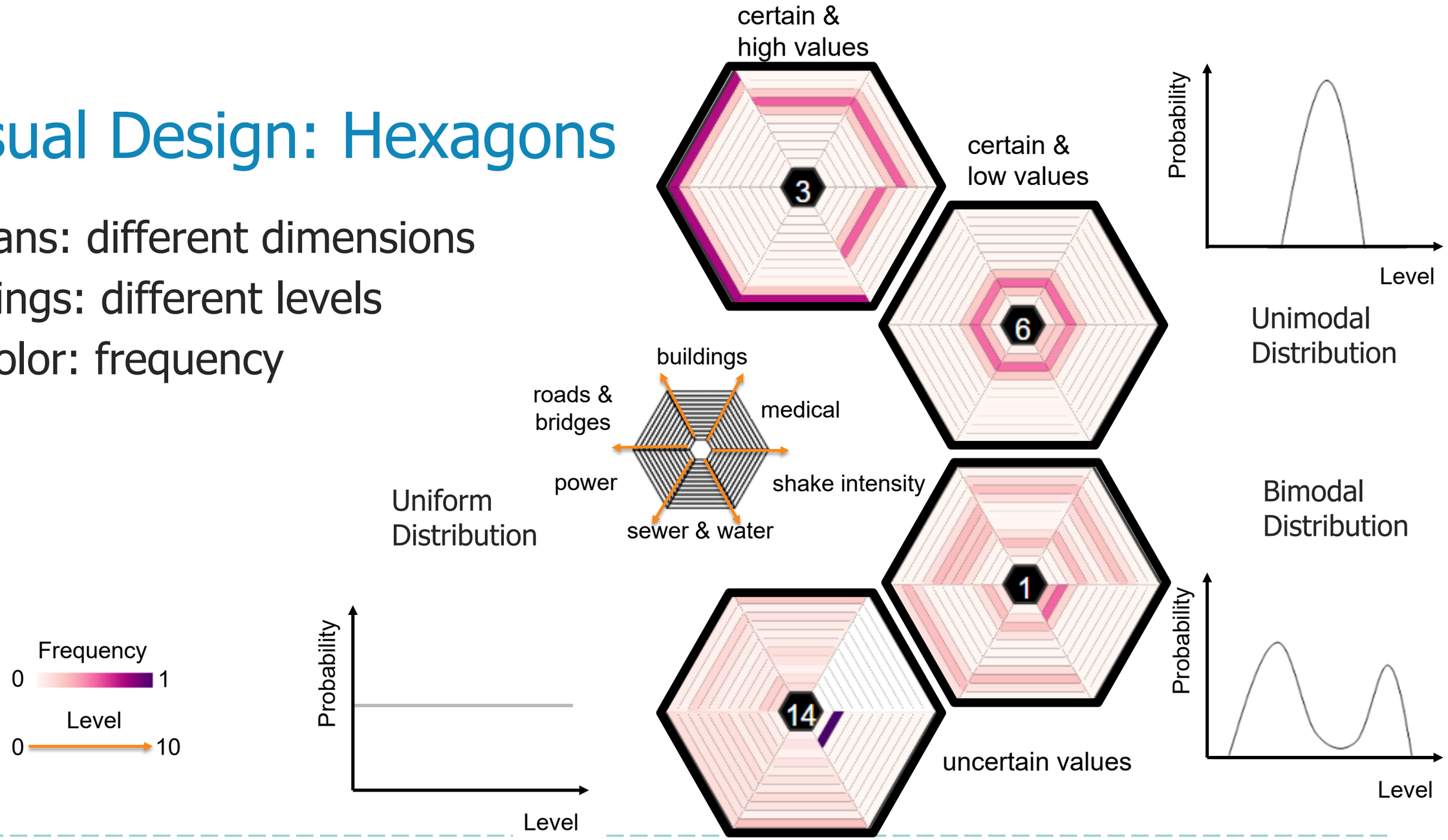
- Hexagon map: encode multivariate data & summarized data





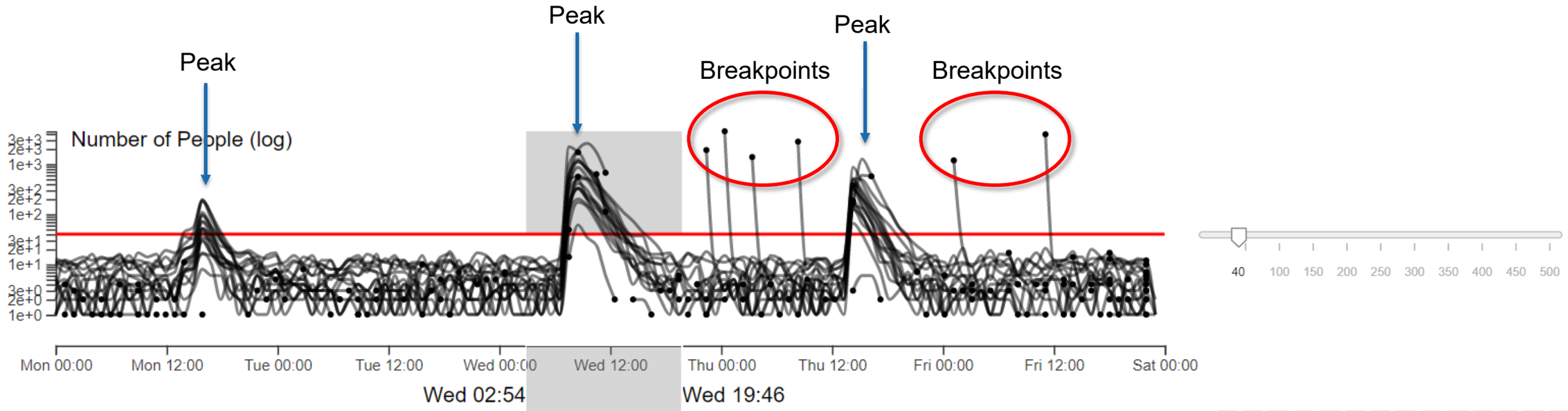
Visual Design: Hexagons

- Pans: different dimensions
- Rings: different levels
- Color: frequency



Visual Design: Line chart

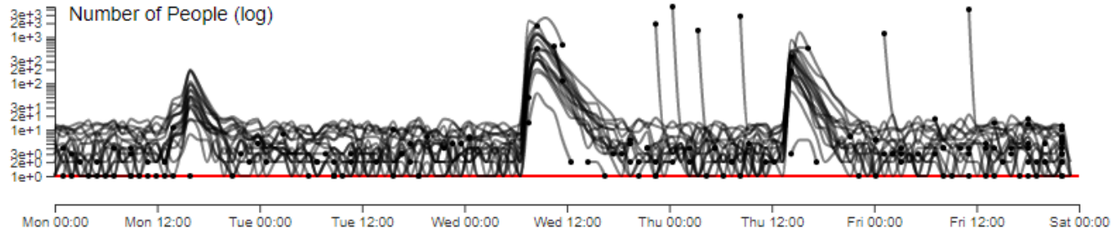
- Number of records every neighborhood along the time
- Filter control panel
- Peaks: pre-earthquake & major earthquakes
- Breakpoints: may due to the power outages

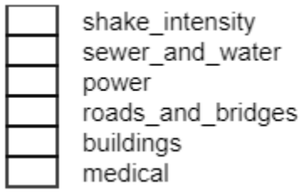


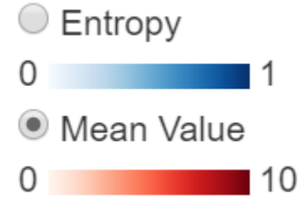
Visual Design: Heat map

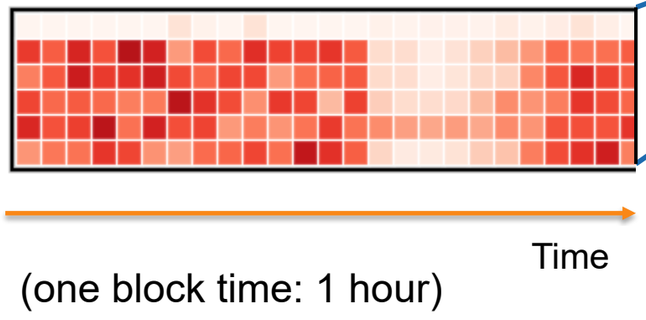
- Measure the damage
- Mean value $S(d_j)$

d_j: dimension j





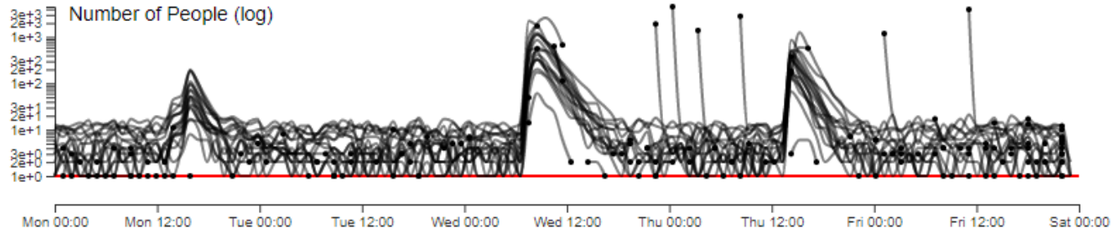


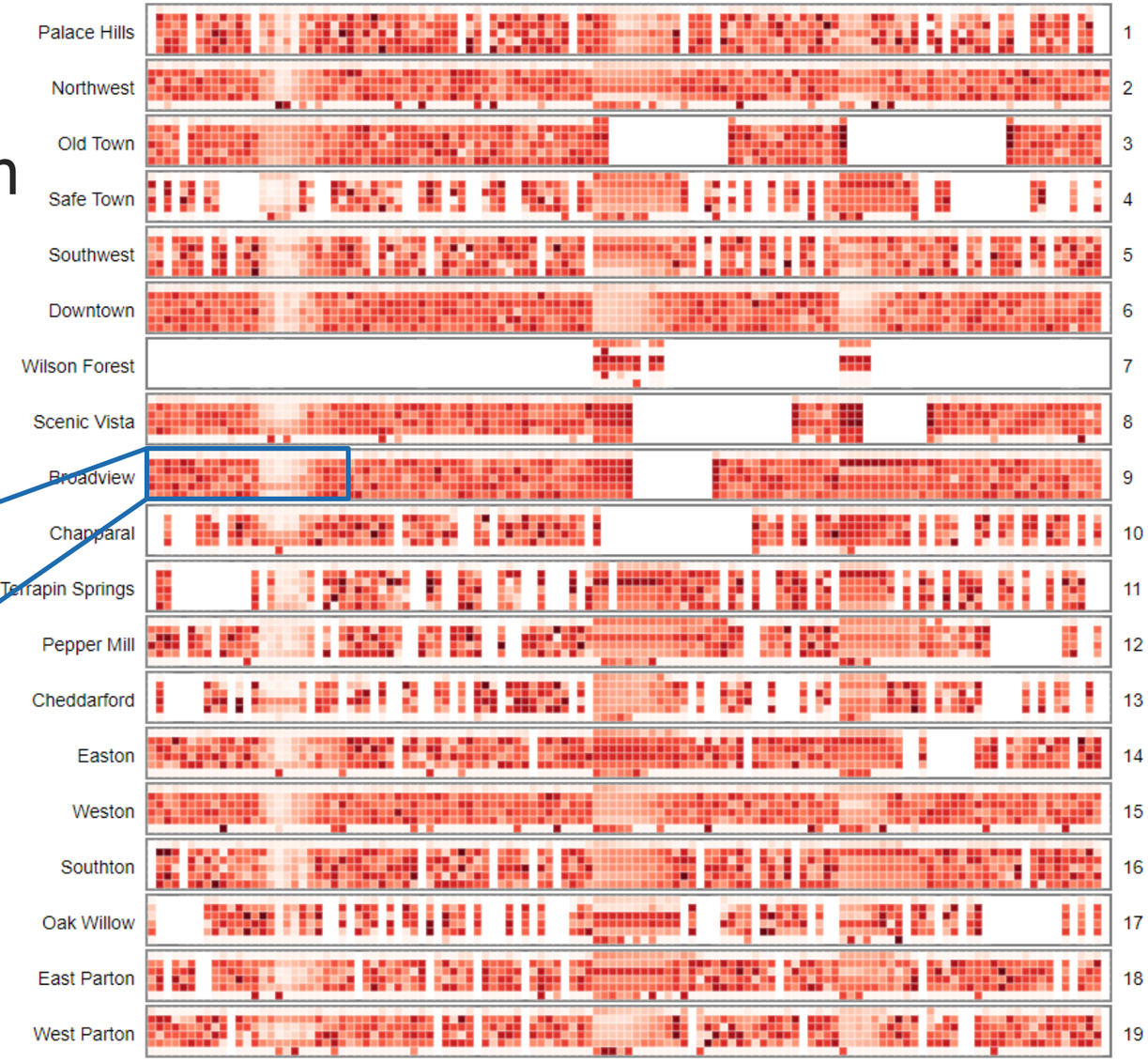


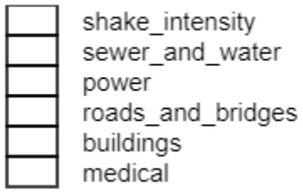
(one block time: 1 hour)

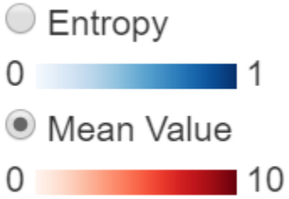
Visual Design: Heat map

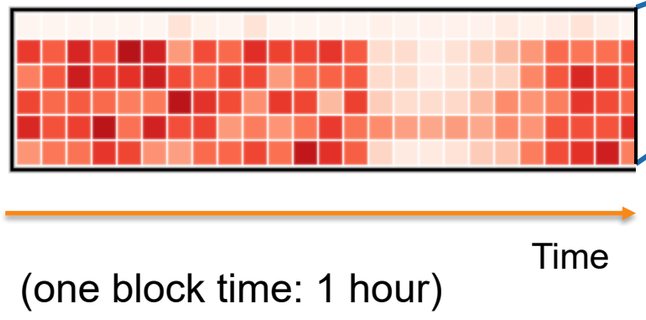
- It seems that the number of the records has negative correlation with the mean values !

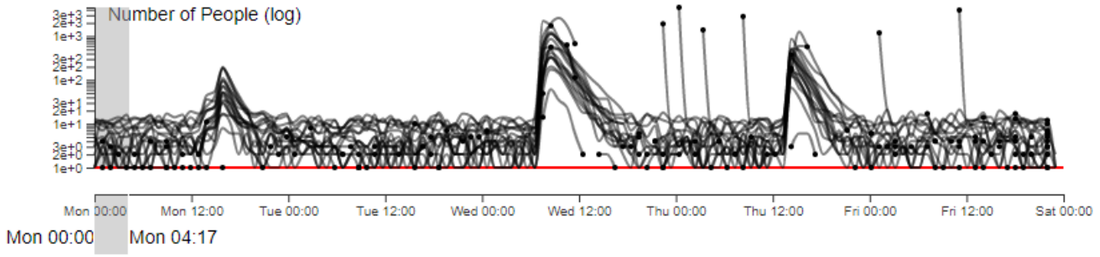


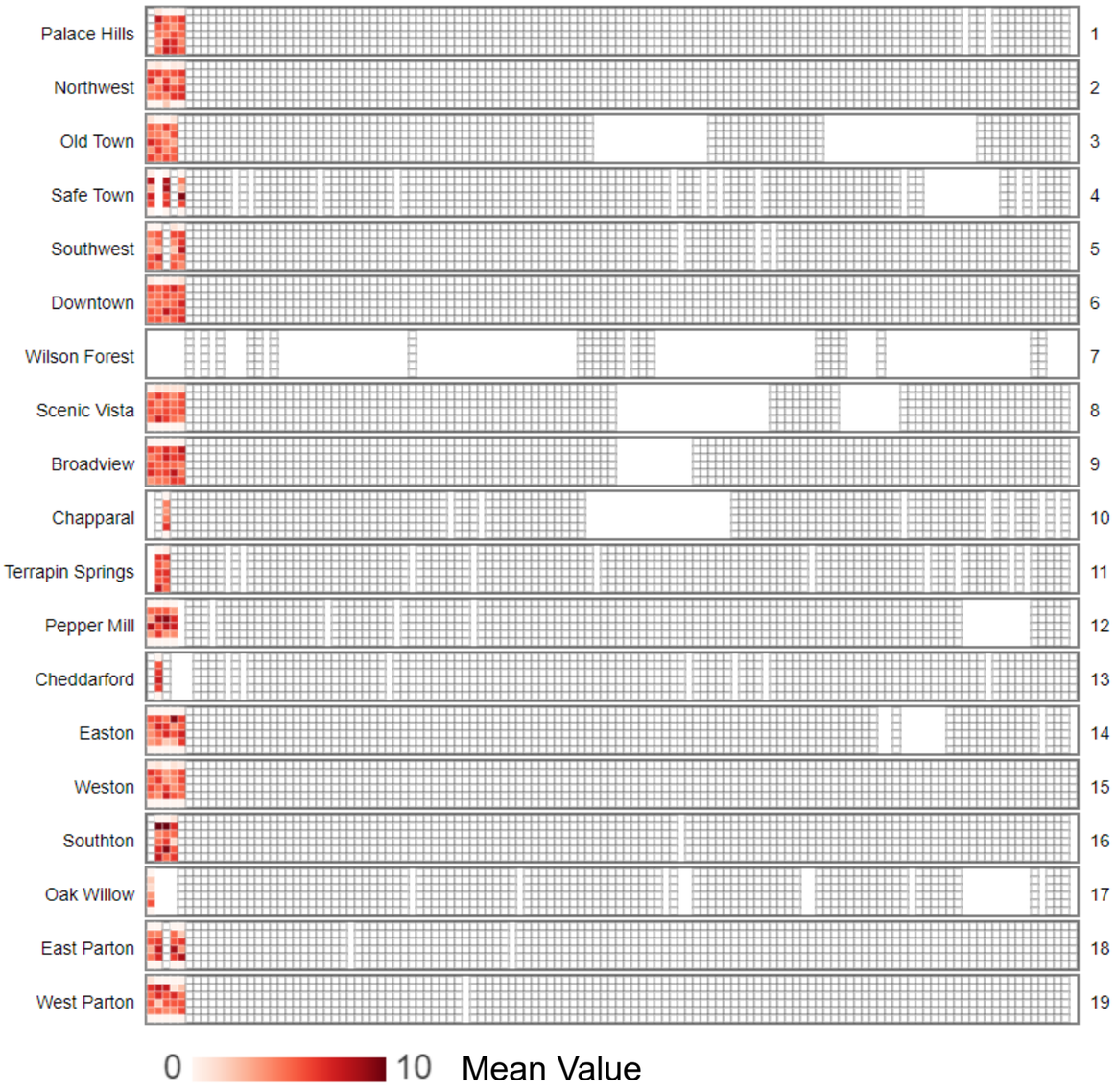




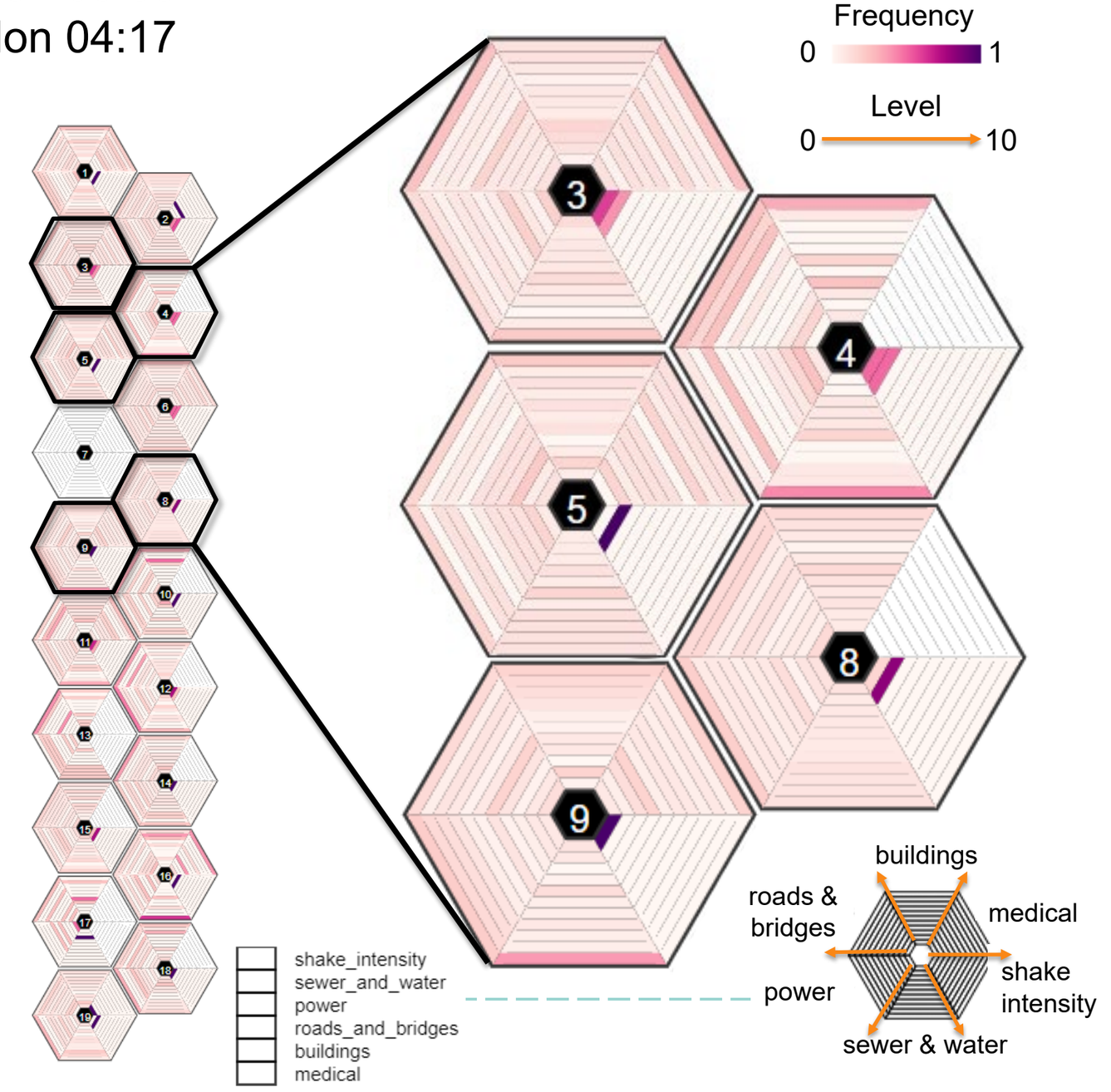


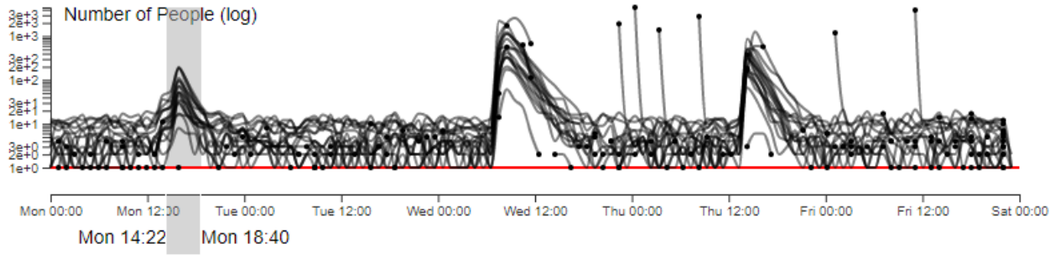




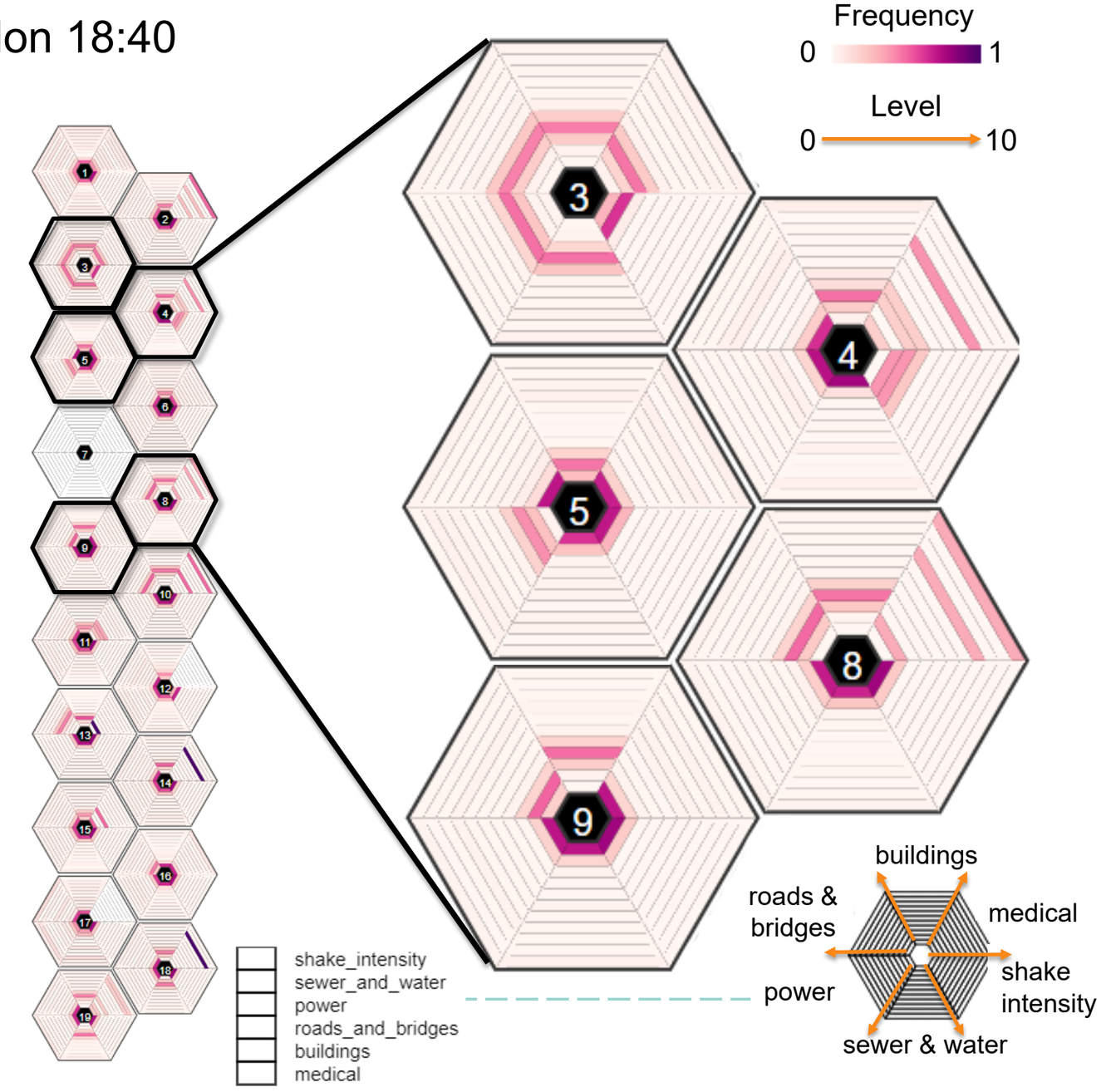


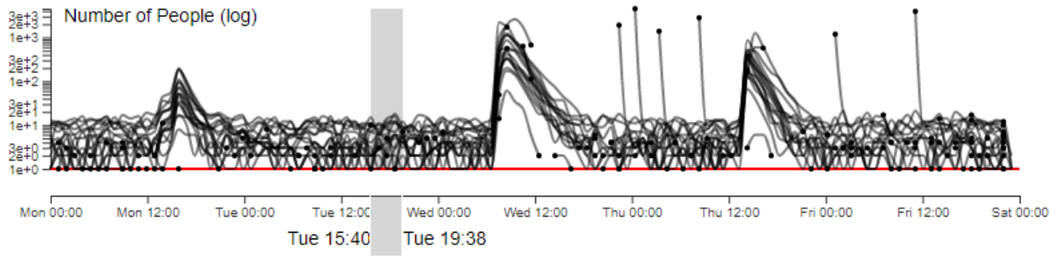
Time:
Mon 00:00 ~
Mon 04:17

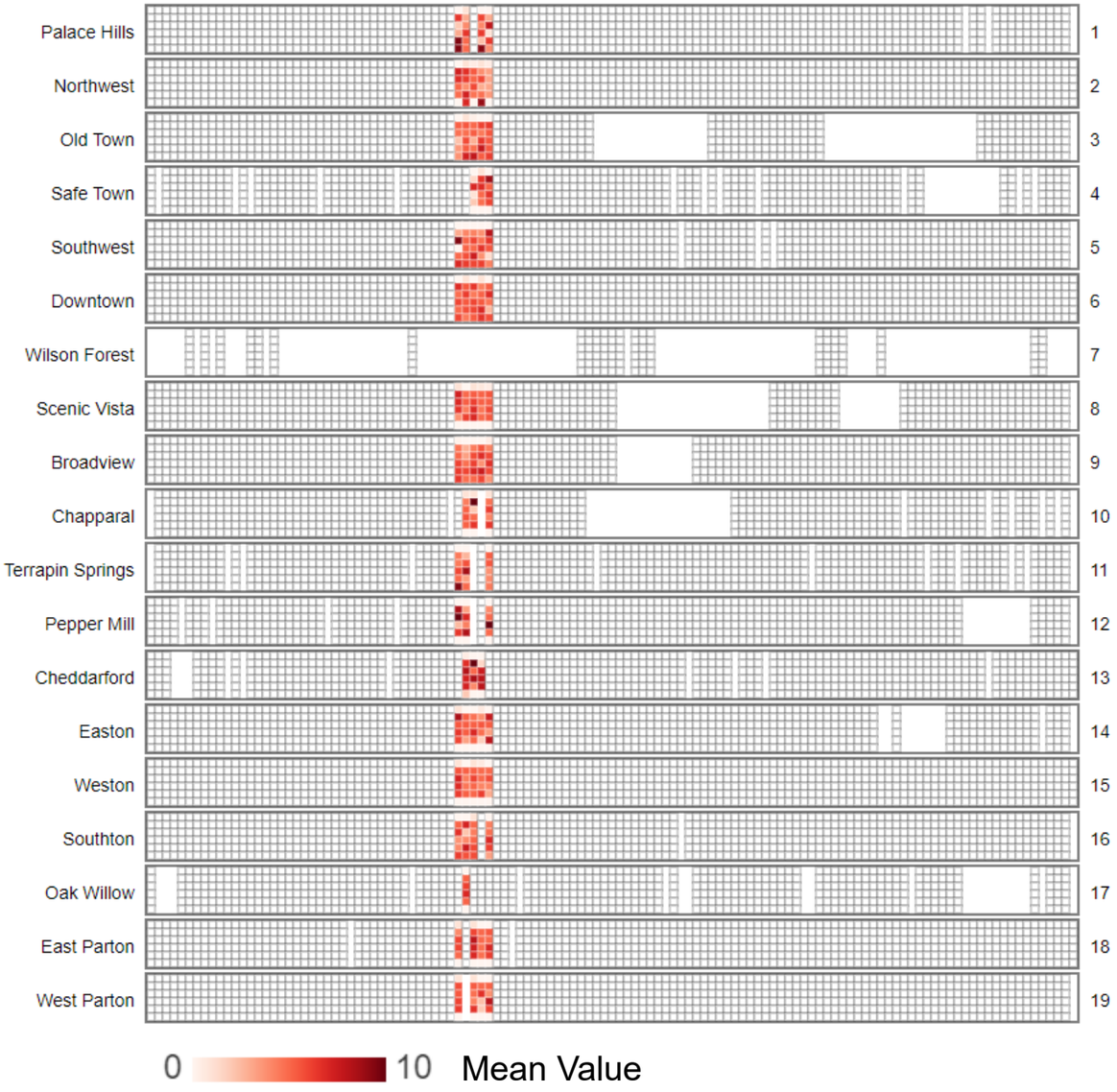




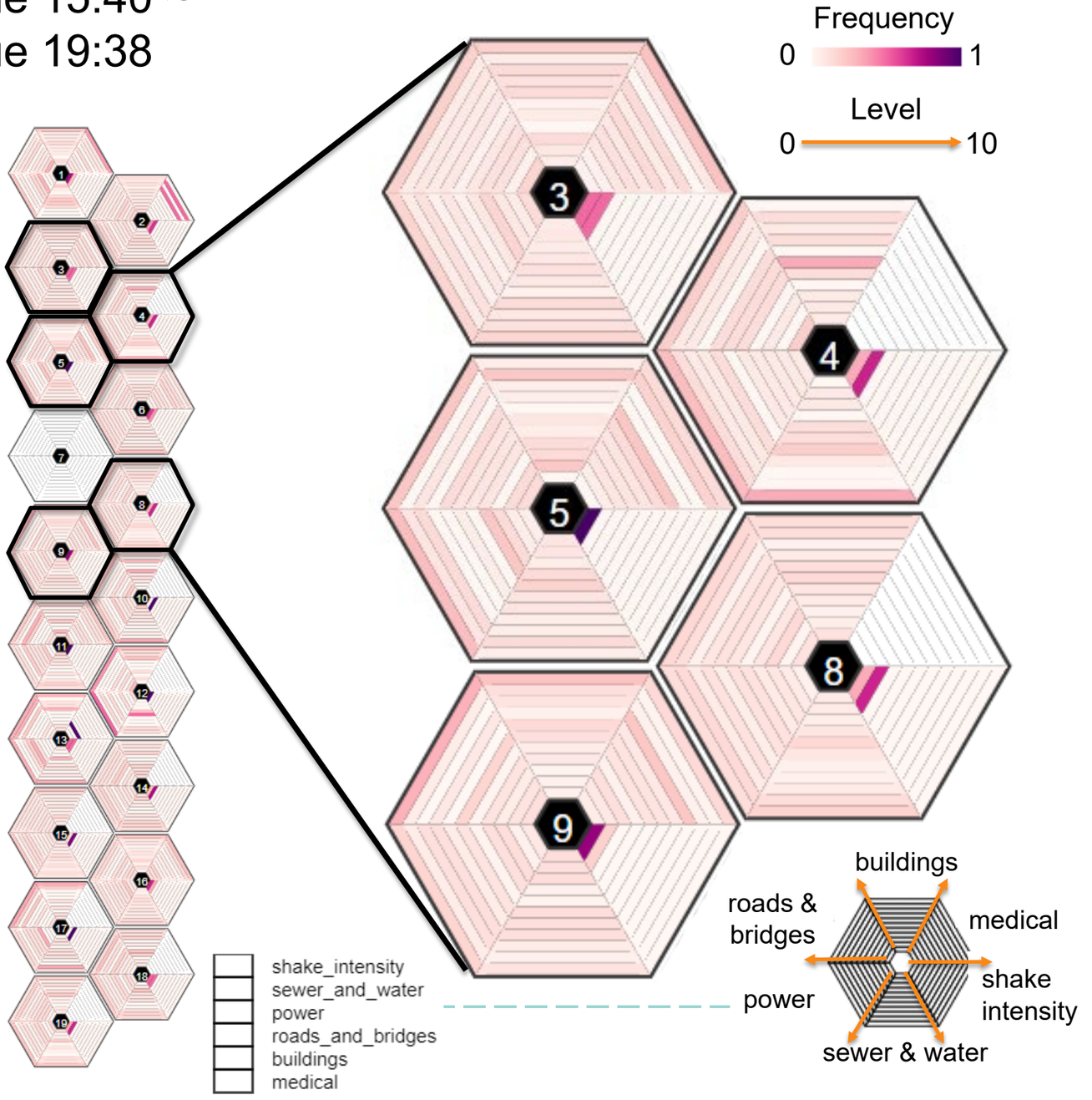
Time:
Mon 14:22 ~
Mon 18:40

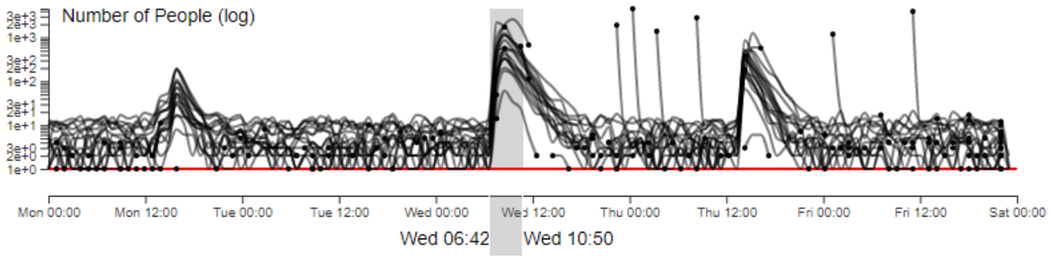




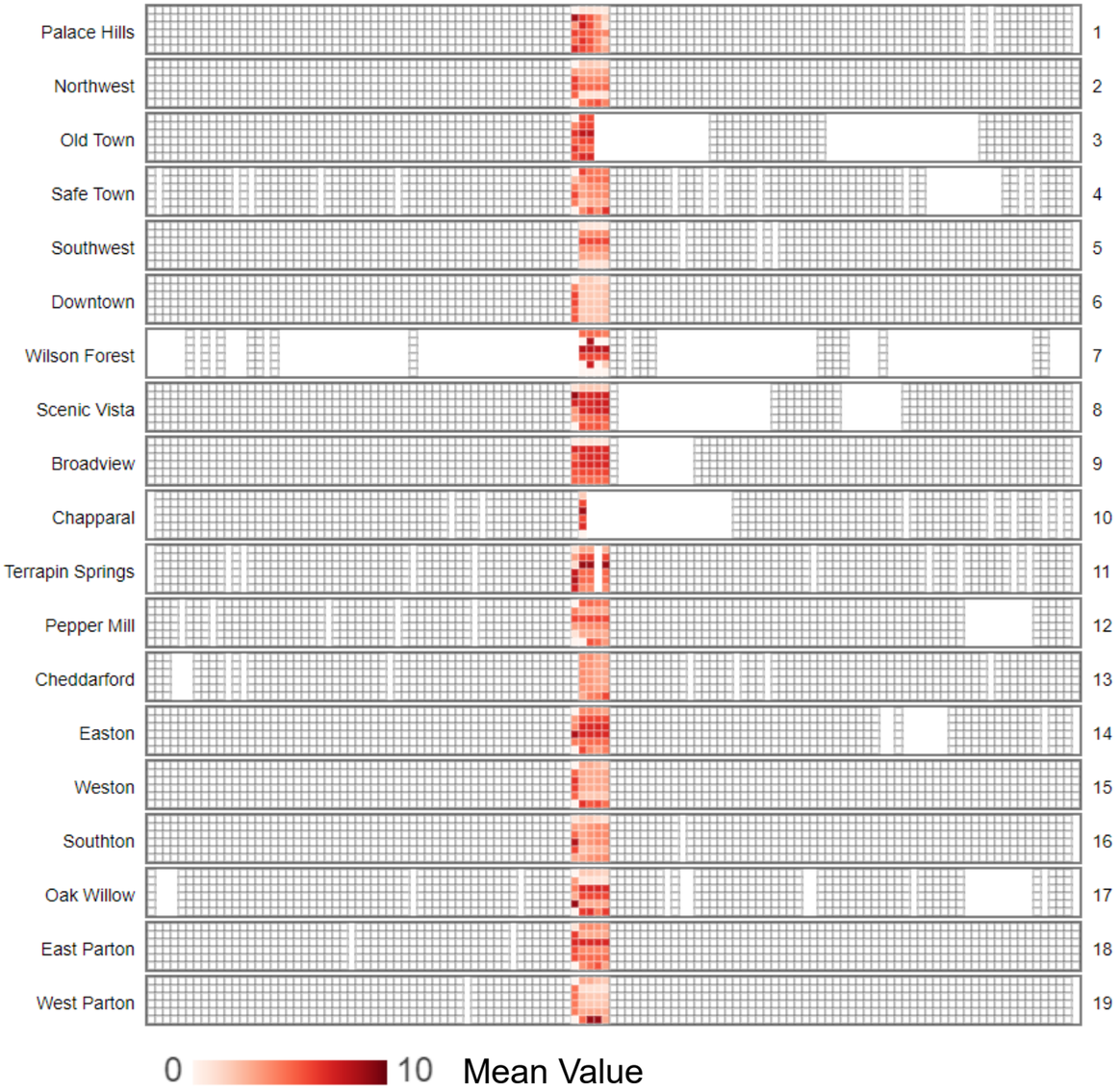


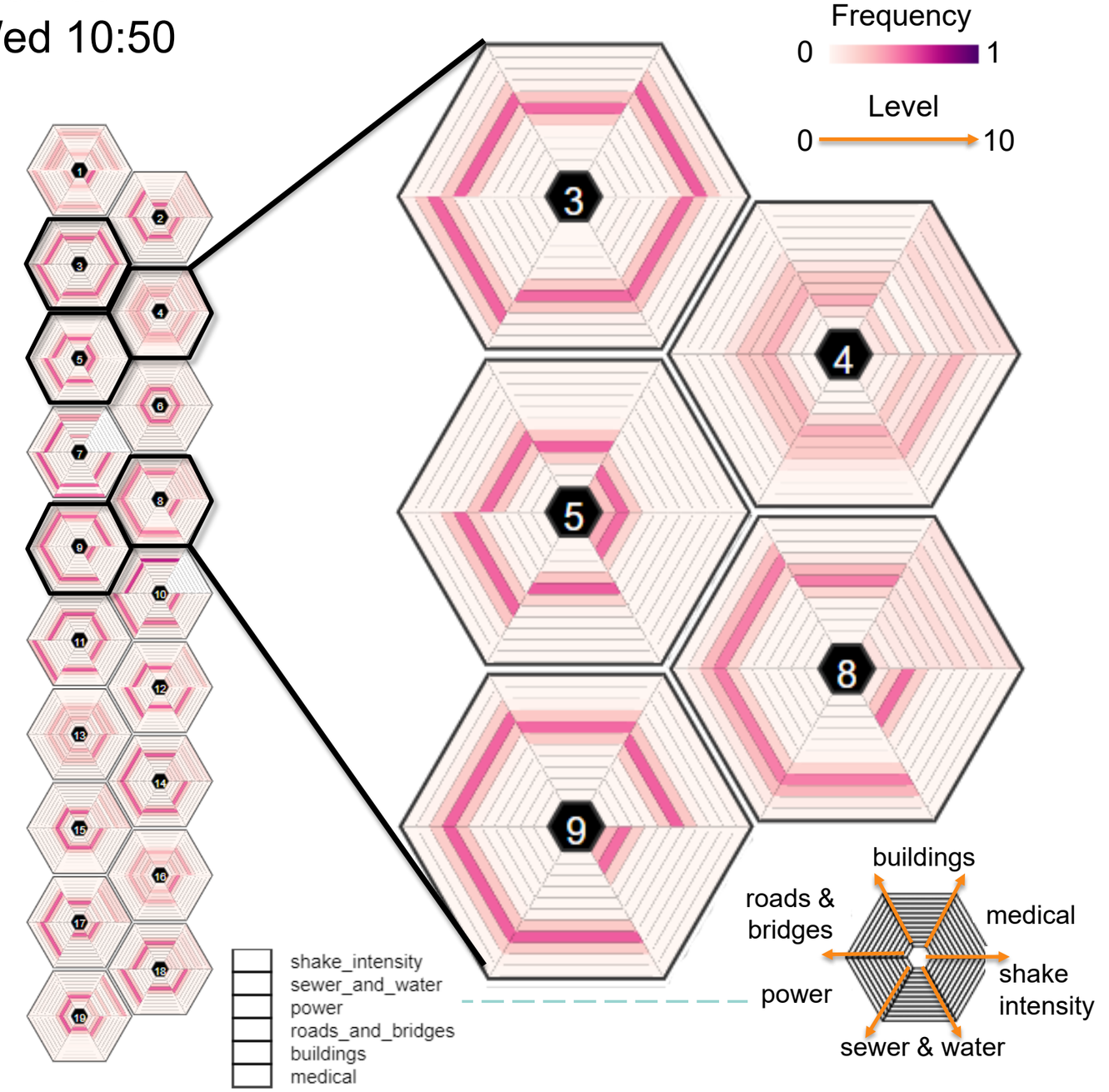
Time:
Tue 15:40 ~
Tue 19:38

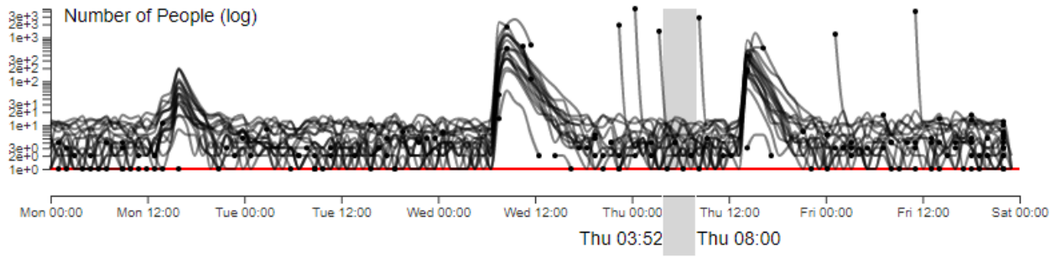




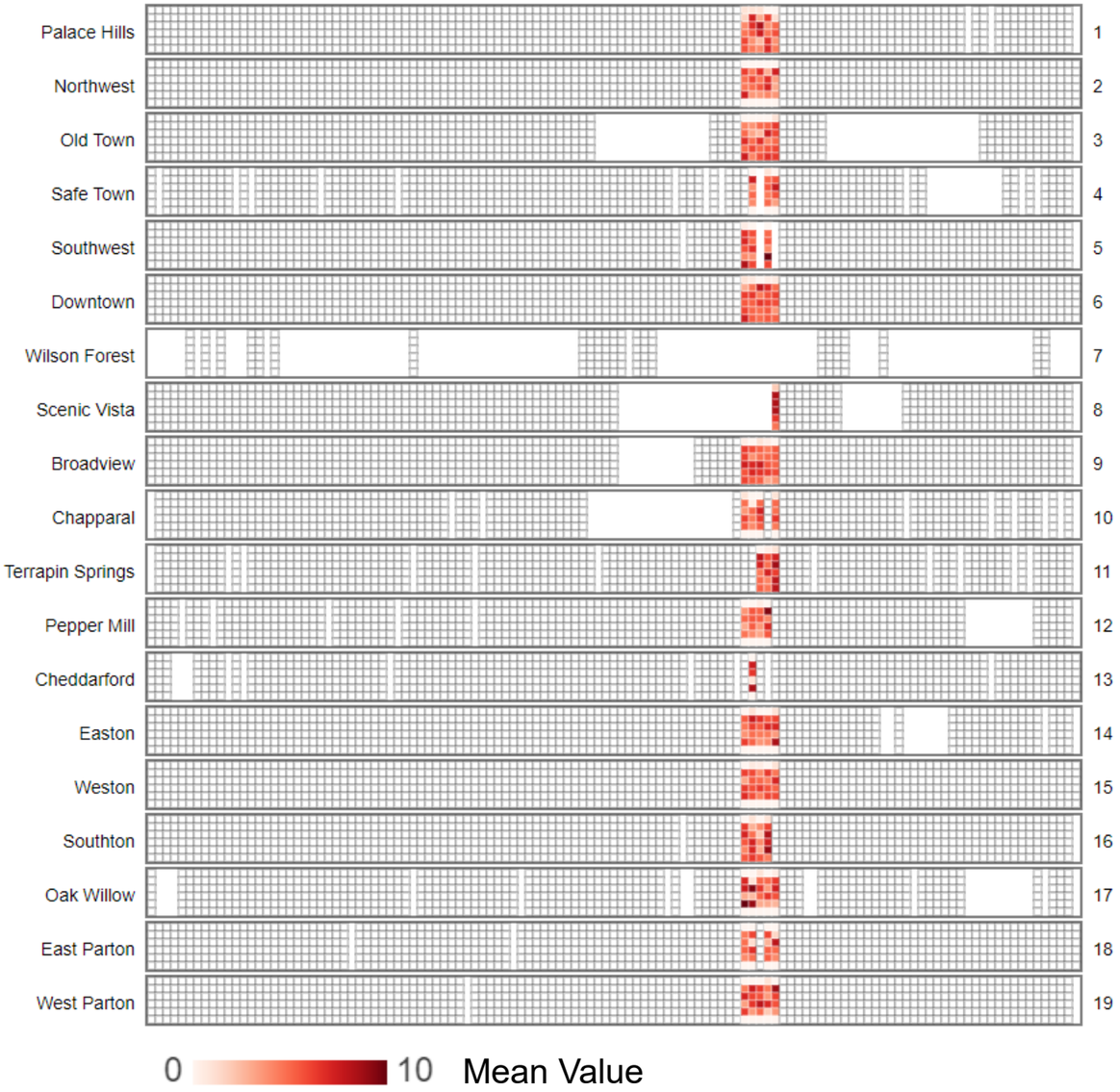
Time:
Wed 06:42 ~
Wed 10:50

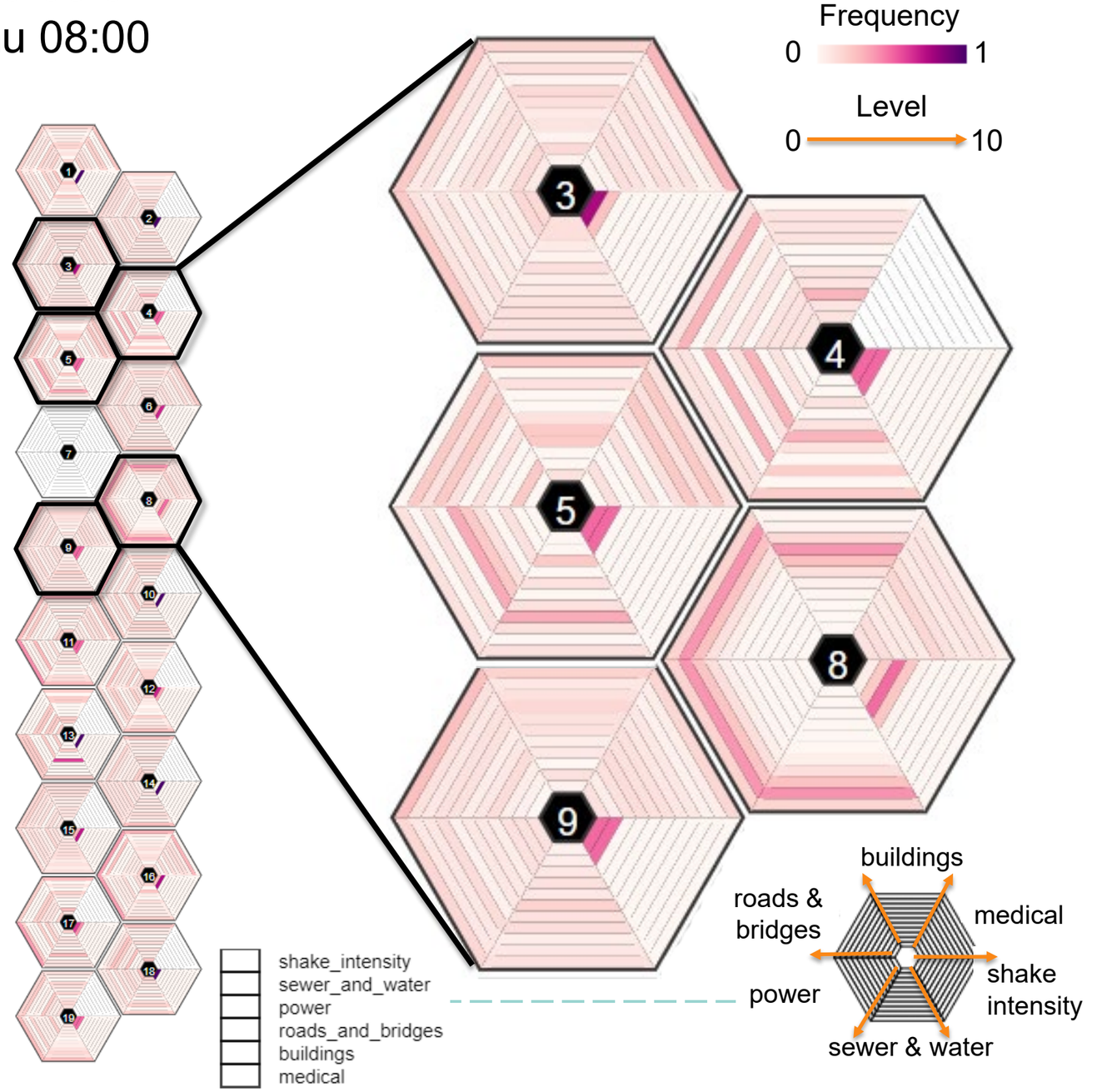


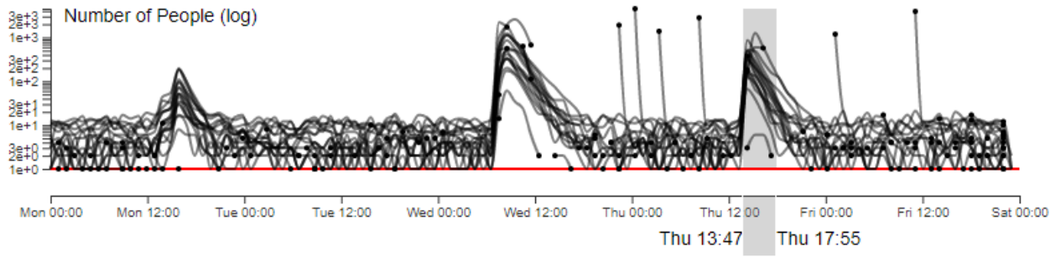




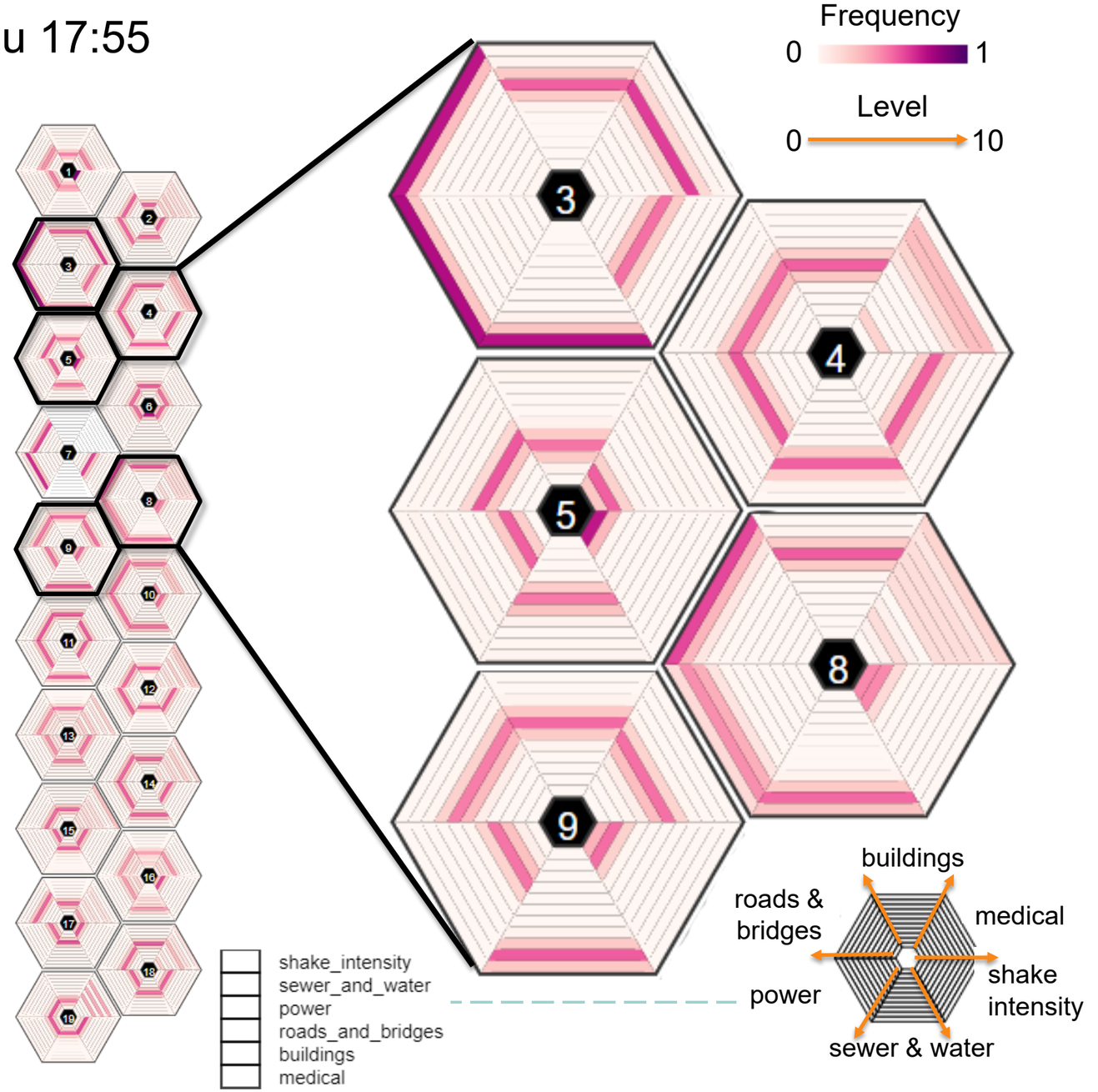
Time:
Thu 03:52 ~
Thu 08:00

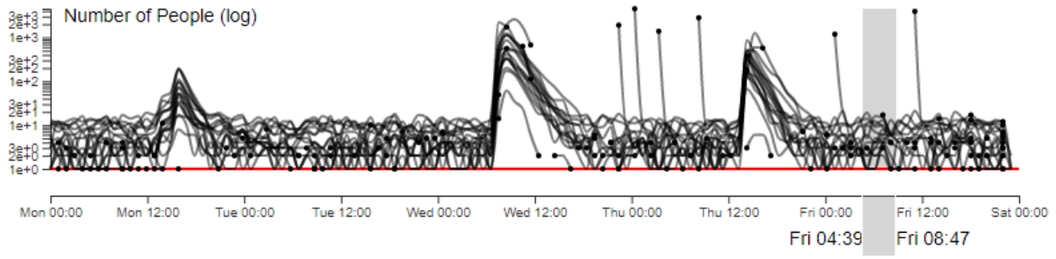




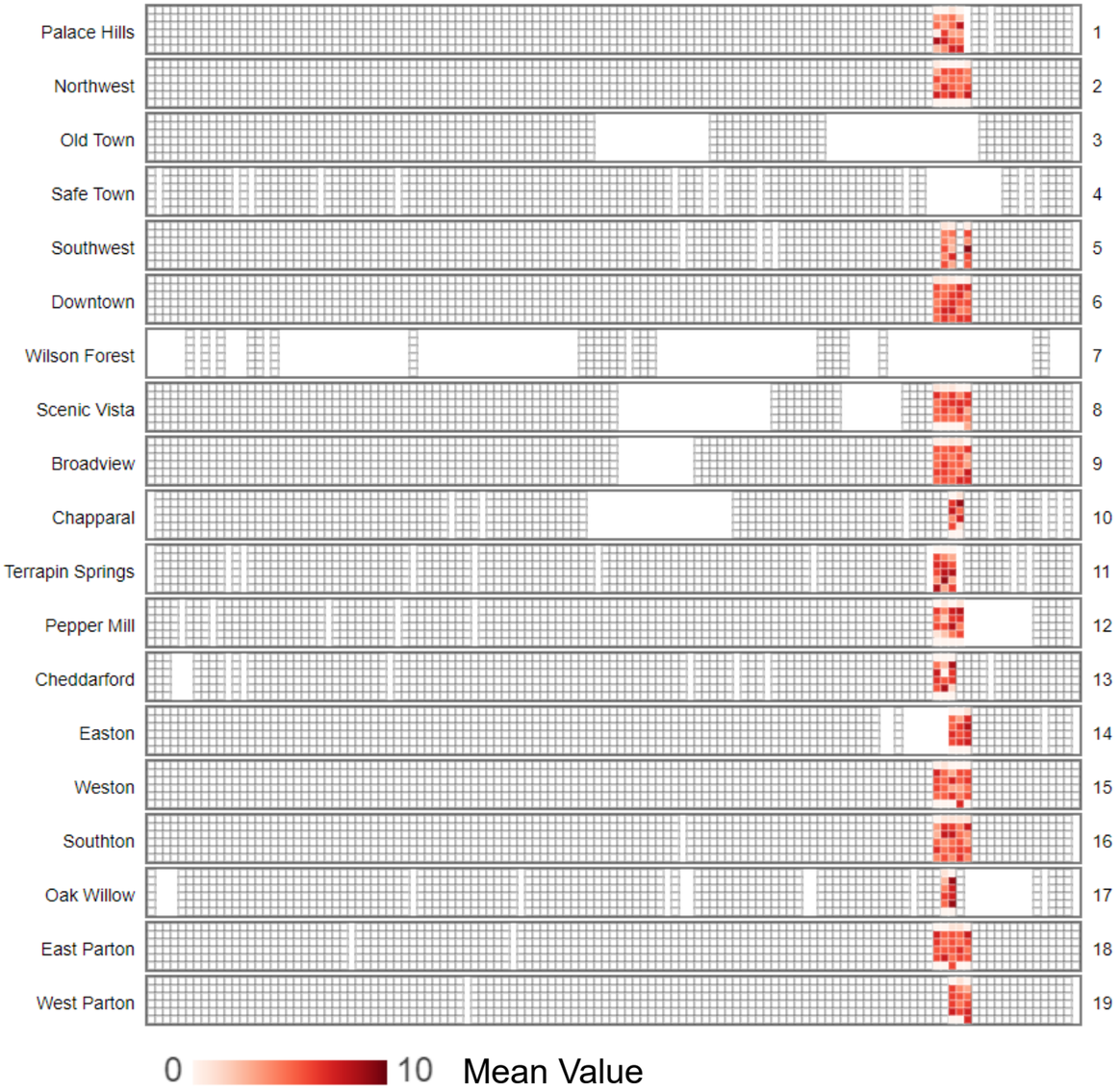


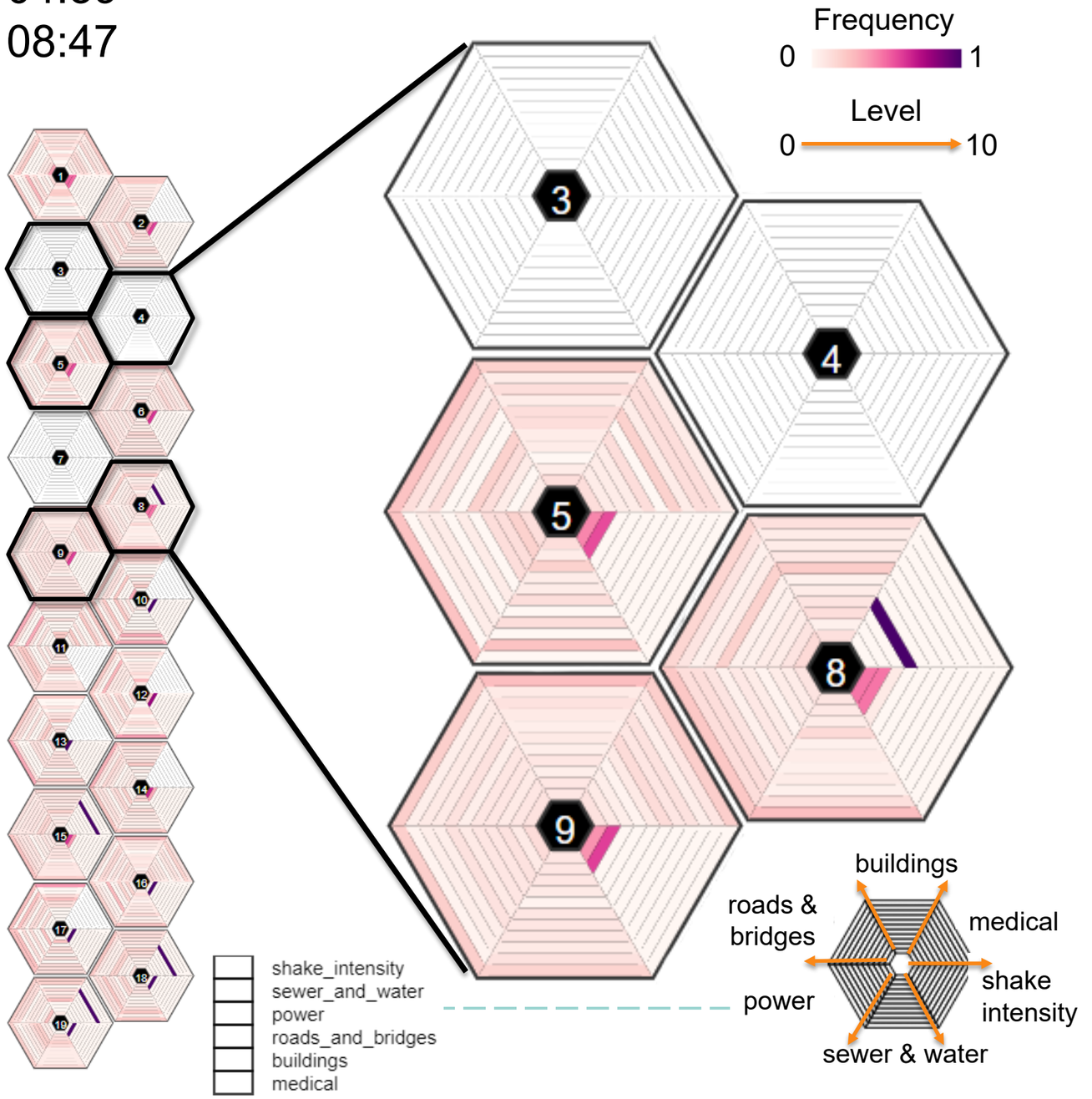
Time:
Thu 13:47 ~
Thu 17:55





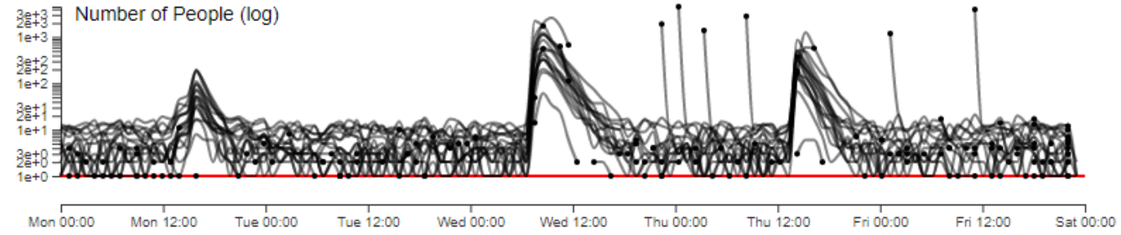
Time:
Fri 04:39 ~
Fri 08:47





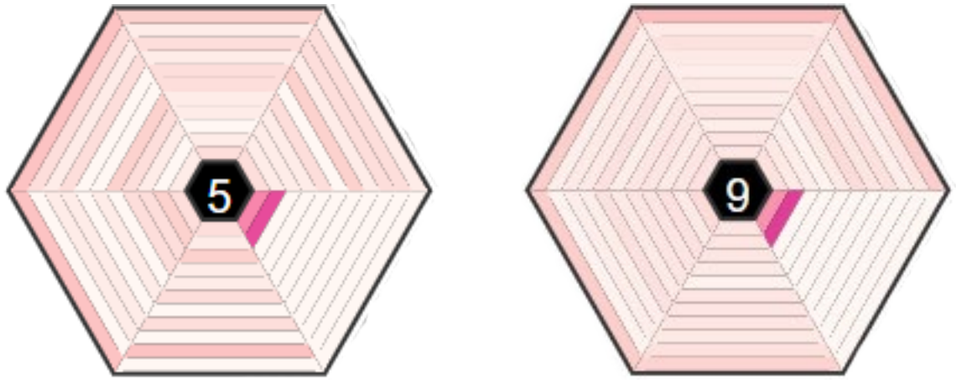
Problem

- Without considering uncertainty, it seems that emergency responders should respond every time, even when earthquakes do not break.



Problem

- It is not easy to compare uncertainty using only hexagons.
- Ex. For hexagons 5 and 9, which record is more uncertain?
- Ex. For hexagon 9, which dimension is more uncertain?



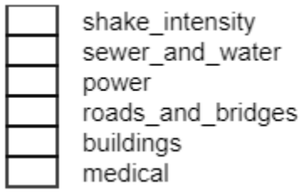
Quantification of uncertainty !

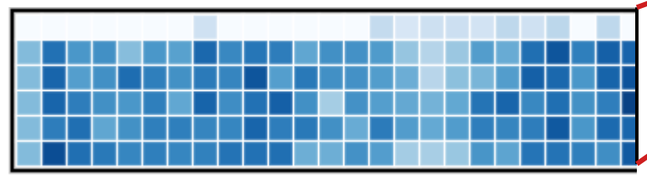
Visual Design: Heat map

- Qualification of uncertainty
- Normalized entropy
- $H_{11}(d_j) = -\sum_{i=0}^{10} p_i(d_j) \log_{11} p_i(d_j)$

  d_j: dimension j, $p_i(d_j)$: frequency of the level i

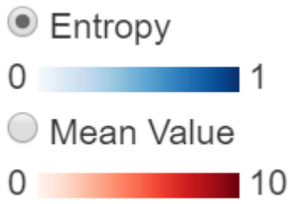
- $H_{11}(d_j) \in [0, 1]$

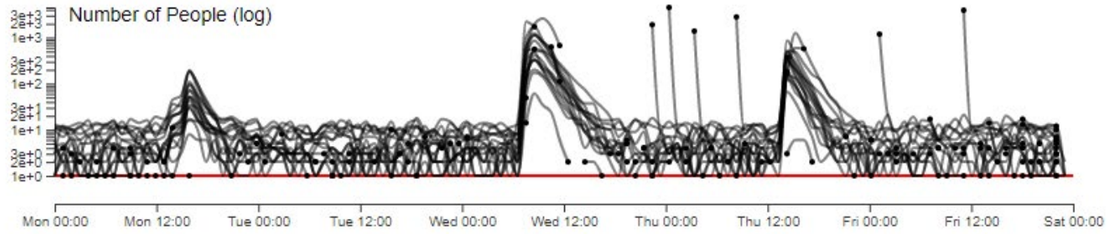




Time

(one block time: 1 hour)

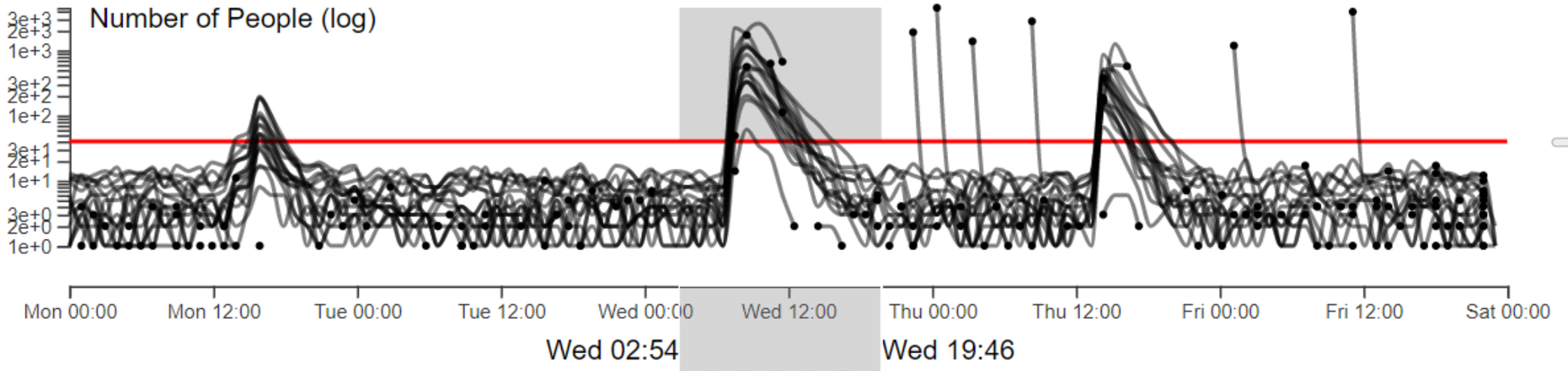




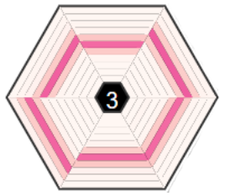
Visual Design: Interactive Rank

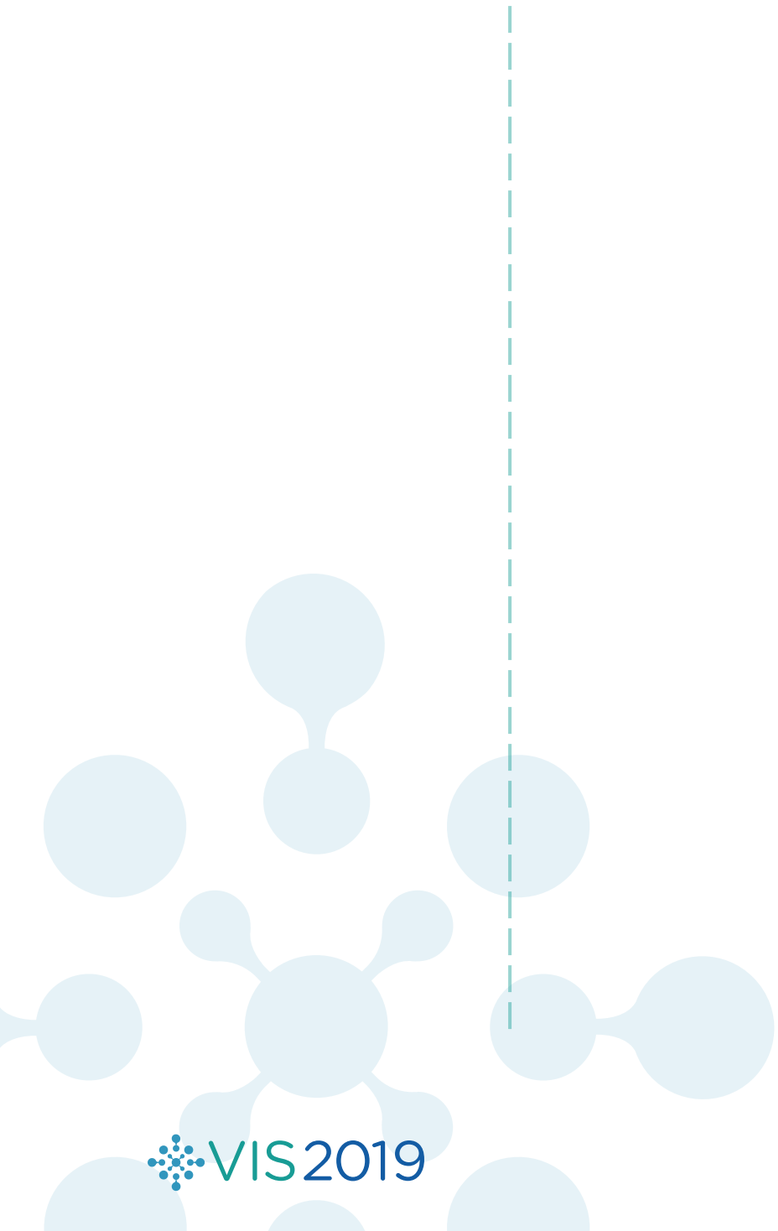
- Rank based on overall damages: $\bar{S} = \frac{\sum_{d_j} \sum_{t_k} S(d_j)/|t_k|}{|d_j|}$
- Rank based on mean entropy: $\bar{H} = \frac{\sum_{d_j} \sum_{t_k} H_{11}(d_j)/|t_k|}{|d_j|}$
- d_j is controlled by :
- t_k is controlled by :

Unselected
Selected



Demo




VIS2019

Conclusion

- Quantification of Uncertainty
- Interaction to rank time series
- Future: can be easily applied to the streaming setting

VIS 2019

Any Questions?

MC1 Award for Excellent Quantification of Abnormalities

Shichao Jia, Jiaqi Wang, Zeyu Li, and Jiawan Zhang
Tianjin University