Voila: Visual Anomaly Detection and Monitoring with Streaming Spatiotemporal Data

Nan Cao, Chaoguang Lin, Quihan Zhu, Yu-Ru Lin, Xian Teng, Xidao Wen.

IEEE Transactions on Visualization and Computer Graphics (Volume: PP, Issue: 99)

http://ieeexplore.ieee.org.ezproxy.library.ubc.ca/stamp/stamp.jsp?arnumber=8022952&tag=1

Presented by: Shirlett Hall
November 28, 2017

Implementation of the map visualization system

High Level Goals:

- To process large scale, dynamic streaming data to detect anomalies
- To allow human inspection and interpretation to guide final machine processes

High Level Features:

- Online Data Processing Pipeline that remains connected to data inputs
- Uses a tensor-based algorithm to produce descriptive patterns over time and space
- Incorporates unsupervised Machine Learning Techniques during human interactions
- Shifts between map modes dependent on user goals

What: Data

- GIS Data in big data scenarios

- Transformed into a sequence of tensor time series at the granular level of an hour, a day, a week, or month

Why: Abstract Tasks

- In a specific point in time, identify spatial locations and objects
- Direct user's attention to potentially significant anomaly instances
- Compare behaviors at the same or different time intervals using user judgement

Why: Domain Tasks

1. Use anomaly detection algorithm against the multi-faceted data
2. Create rich-context visualizations that show suspicious patterns from the tensor analysis
   - Overview -> ranking -> link to raw data
   - Showing patterns -> comparing patterns -> external memorization
3. Apply or update Bayesian rules as users re-order anomalous patterns by degree of importance

Why: Domain Tasks, cont'd

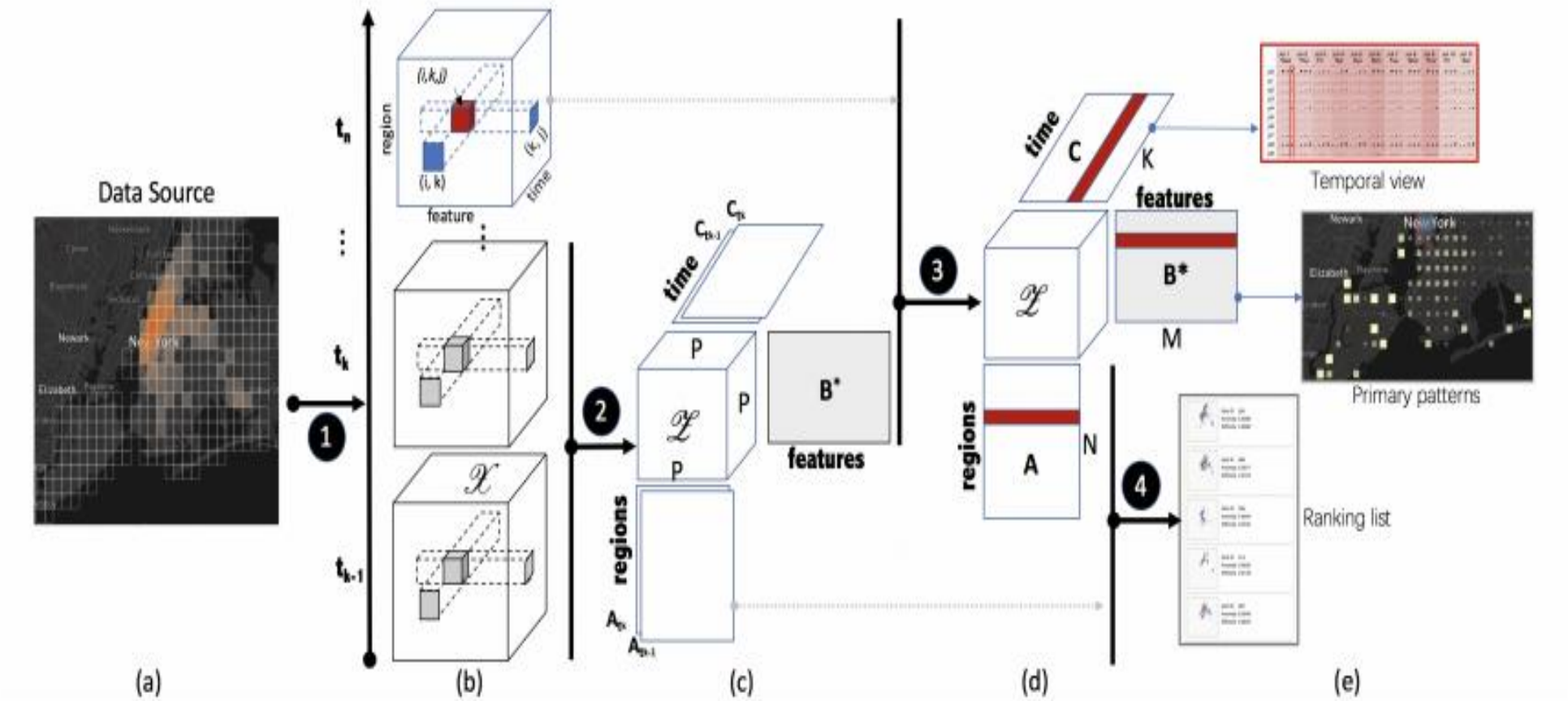


Fig. 4. The visual anomaly detection in the streaming spatiotemporal data consists of four major steps: (1) transforming the streaming spatiotemporal data into a tensor time series, (2) the expected pattern analysis based on historical data, (3) context analysis based on tensor decomposition, and (4) online regional anomaly detection in context.

How: Facets

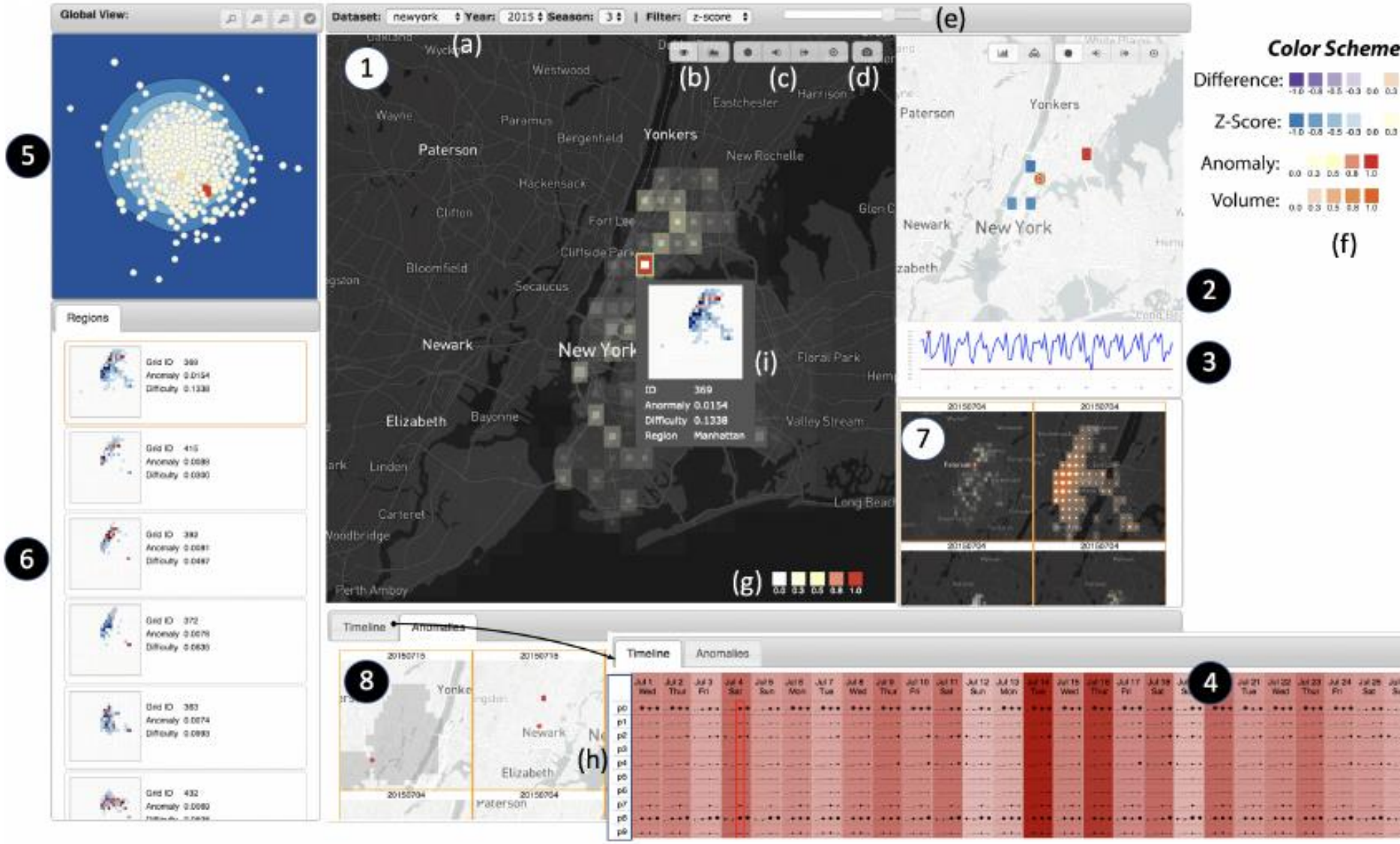


Fig. 5. The user interface of Voila system consists of eight major views: 1) macro and 2) micro map views; 3) history view; 4) te
5) feature inspection view; 6) ranking list; 7) snapshot panel; and 8) anomaly panel.

How: Heatmap in Anomaly Detection Mode

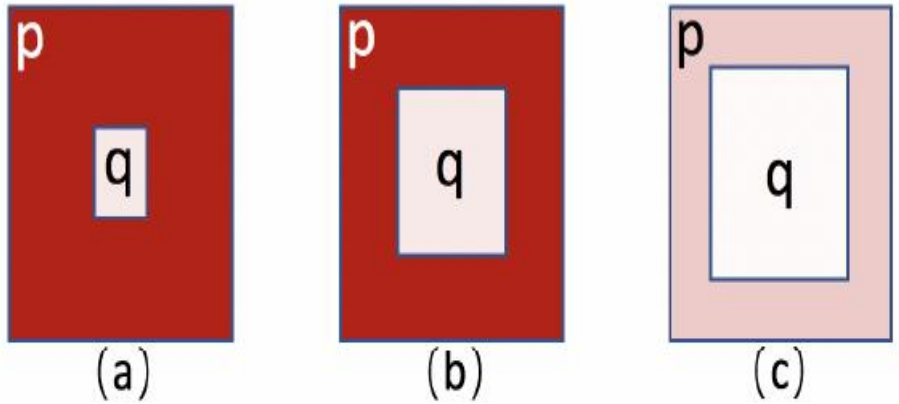
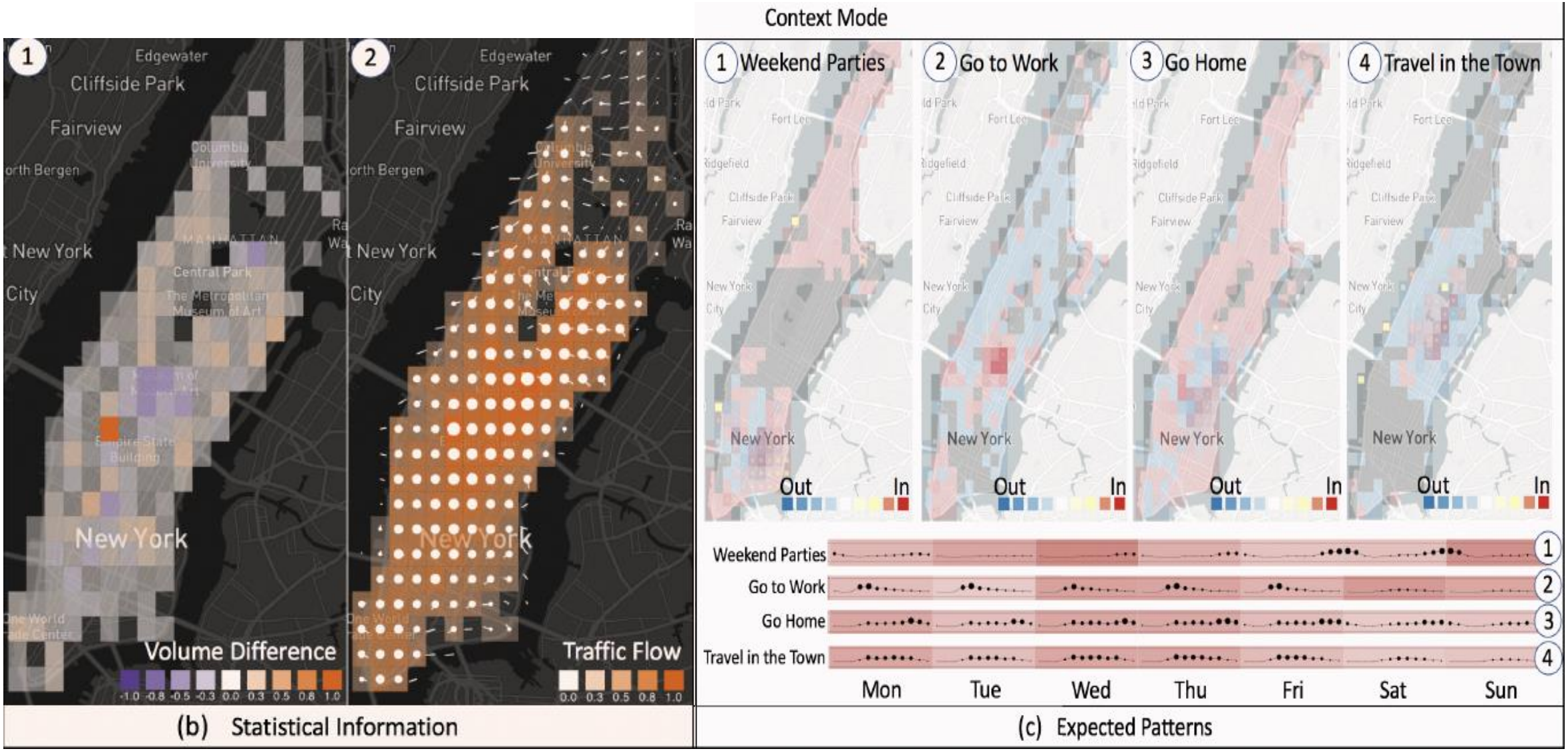


Fig. 6. The anomaly glyphs reveal both anomaly likelihood (saturation of the background color) and the difficulty of finding an anomaly (the size of the inner rectangle).

How: Heatmap in Context Mode



Voila - Short Video showing Application

Voila : Visual Anomaly Detection and Monitoring with Streaming Spatiotemporal Data

Submission ID: # 230

Analysis Summary

System	Voila
What:Data	Streaming, dynamic GIS with spatial location and objects
What: Derived	Tensor with features, space and time as quantitative attributes
Why: Tasks	Apply anomaly detection, show suspicious patterns, compare historical patterns, allow analysis, process user feedback, update machine learning algorithm
How: Encode	Dense spatial area using rectangular layered glyphs; colormaps with diverging hues and sequential saturation levels; and popouts/tooltips on fields
How: Facet	Multiform linked layouts including 2 views with a main map and a less detailed map to show context; time series showing area history; tabular chart showing anomalies; panel showing ranking of multiple regions
How: Reduce	Filtering
How: Manipulate	Selection and highlighting, pan, zoom, brush
Scale:	up to 311 grids on a map, over 100 million instances, ex. volume of traffic as attribute

Evaluation and Next Steps: Authors' Perspective

- The tensor detection method produced more satisfactory positive identification rates than other baseline methods
- With the aid of the system's visual tools, users are well prepared to fixate on only suspicious events
- Since the initial visualization seems overwhelming at first glance, need tutorials
- More visual clues
- Support for Fact Checking
- Adaptively determine granularity with respect to time
- Embed forecasting and prediction capability

Additional Critique

- Good that they have included a human in loop
- Channels are noticeable and fairly effective for the intended purpose with some exceptions, ex. rainbow like color map for z-scores
- The purpose of each juxtaposed view is not clear to a novice user
- The authors should have more than one domain expert provide feedback
- When a particular anomaly is noted as being normal, then this may increase the likelihood that false negatives occur in the future
- Size of the q inside the glyph, should be a number
- No mention of the system requirements to process and store so much data and the speed of the algorithm