Voila: Visual Anomaly Detection and Monitoring with Streaming Spatiotemporal Data

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

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

Analysis Summary
<table>
<thead>
<tr>
<th>System</th>
<th>Voila</th>
</tr>
</thead>
<tbody>
<tr>
<td>What: Data</td>
<td>GIS Data in big data scenarios</td>
</tr>
<tr>
<td>What: Derived</td>
<td>Transformed into a sequence of tensor time series at the granular level of an hour, a day, a week, or month</td>
</tr>
</tbody>
</table>

How: Heatmap in Anomaly Detection Mode
- 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

How: Facets
- Deep spatial area using rectangular layered glyphs; colormaps with diverging hues and sequential saturation levels; and popouts/tooltips on hover
- Multiform linked layouts including 2 views with a main map and a less detailed map to show context; time series showing area history; tabular charts showing anomalies; panel showing ranking of multiple regions
- Selection and highlighting, pan, zoom, brush up to 311 grids on a map, over 100 million instances, ex. volume of traffic as attribute
- Filtering

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 inside the system’s requirements to process and store so much data and the speed of the algorithm
- 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
- 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

Why: Domain Tasks, cont’d
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
4. In a specific point in time, identify spatial locations and objects
5. Direct user’s attention to potentially significant anomaly instances
6. Compare behaviors at the same or different time intervals using user judgement

How: Heatmap in Context Mode
- 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

How: Short Video showing Application
- Voila - Short Video showing Application

Voila: Visual Anomaly Detection and Monitoring with Streaming Spatiotemporal Data

Submission ID: 250