VAST Challenge 2017 Reviewer Guide: Mini-Challenge 2

This document provides information to support peer review of submissions to VAST Challenge 2017, Mini-Challenge 2. This document covers background about the submission structure, the challenge problem, tasks and questions presented to participants, potential answers, and evidence found in the Challenge data supporting these answers. For a full description of the challenge problems and to access the data provided to the participants, please visit http://vacommunity.org/VAST+Challenge+2017.

Submissions

Participants are required to submit their entries on a standard answer form, along with a video explaining how visual analytics were used to help solve the challenges. Please consider both parts of the submission in your review. If you have difficulty reading the answer form or playing the video, please contact us at vast-challenge@ieee.org for assistance.

Scenario

Overview

Mistford is a mid-size city located to the southwest of a large nature preserve. The city has a small industrial area with four light-manufacturing endeavors. Mitch Vogel is a post-doc student studying ornithology at Mistford College and has been discovering signs that the number of nesting pairs of the Rose-Crested Blue Pipit, a popular local bird due to its attractive plumage and pleasant songs, is decreasing! The decrease is sufficiently significant that the Pangera Ornithology Conservation Society is sponsoring Mitch to undertake additional studies to identify the possible reasons. Mitch is gaining access to several datasets that may help him in his work, and he has asked you (and your colleagues) as experts in visual analytics to help him analyze these datasets.

Mini-Challenge 2

Ornithology student Mitch Vogel was immediately suspicious of the noxious gases pouring out of the smokestacks from the four manufacturing factories south of the nature Preserve. He was almost certain that all of these companies are contributing to the downfall of the poor Rose-Crested Blue Pipit bird. But when he talked to company representatives and workers, they all seem to be nice people and actually pretty respectful of the environment.

In fact, Mitch was surprised to learn that the factories had recently taken steps to make their processes more environmentally friendly, even though it raised their cost of production. Mitch discovered that the state government has been monitoring the gaseous effluents from the factories through a set of

sensors, distributed around the factories, and set between the smokestacks, the city of Mistford and the nature preserve. The state has given Mitch access to their air sampling data, meteorological data, and locations map.

The primary job for our intrepid graduate student Mitch, is to determine which (if any) of the factories may be contributing to the problems of the Rose-Crested Blue Pipit. Often, air sampling analysis deals with a single chemical being emitted by a single factory. In this case, though, there are four factories, potentially each emitting four chemicals, being monitored by nine different sensors. Further, some chemicals being emitted are more hazardous than others. The task is to detangle the data to help Mitch determine where problems may be.

It would be helpful for the reviewer to look at the background information on monitoring and the factories. Mini-Challenge 2 has several opportunities for innovative visualizations. As an instructional overview, Figure 1a shows the area of the factories and the sensors, plus a compass for wind direction and a simplified plume, emanating from Factory 2 (Kasios). This image shows that if Factory 2 is emitting air particles, then it is most likely that Sensor 7 will be the one that picks up that particular chemical during that time period. The details of plume dispersion for the chemicals and the environment was not described in detail, so we look to the contestants to make some reasonable assumptions about those elements. There could be some situations where two factories are emitting pollutants that are picked up by the same sensor (Figure 1b). The task of the contestant is to look at other times of emissions, when the wind is not blowing in a direction that would cause and overlap to understand a single factories emission signatures.

In Figure 1c we see a situation where two factories are emitting pollutants, but only one is being read by a sensor. In this situation, the emissions by Factory 4 (Indigo) will not be detected by any of the sensors, due to the position of the factory and the positions of the sensors.

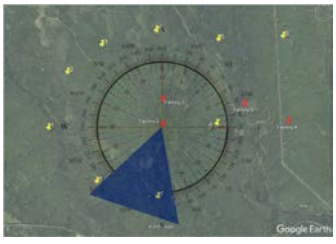

Figure 1a. Single Factory, single sensor reading

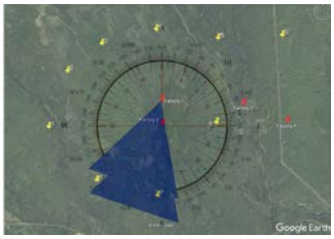

Figure 1b. Two Factories, overlapping sensor readings

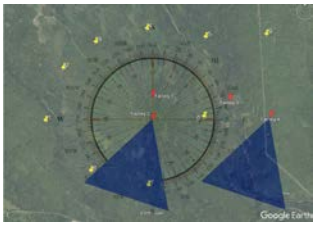

Figure 1c. Two Factories, readings from only one Factory possible

Ground Truth

Kasios (Factory 2) is venting their shop directly to the atmosphere, which releases Methylosmolene. Although there is no regulatory limit established in the challenge, Kasios had claimed to have switched over to the safer AGOC-3A. This has implications for Grand Challenge contestants as the paint booths and guns are cleaned during the evening shift and the barreled sludge needs to be dumped.

Data

The contestants are provided a background description of monitoring around the industries. This is a key piece of information as it identifies Methylosmolene as the worst volatile organic compound (VOC) being monitored. The companies' background information provides descriptions of each of the companies with air emissions. The map is a bitmap representation of the area of interest near the Preserve. The data descriptions explain the factory locations and the sensor locations in terms of the gridded map layout.

The sensor data provides fields for the chemical name, the monitor, the date-time, and the monitor reading. The meteorological data provides a time stamp, wind direction and wind speed. The data is provided for 3 months across the scenario timeline: April, August, and December of 2016.

Contestant Questions

1. *Characterize the sensors' performance and operation. Are they all working properly at all times? Can you detect any unexpected behaviors of the sensors through analyzing the readings they capture? Limit your response to no more than 9 images and 1000 words.*

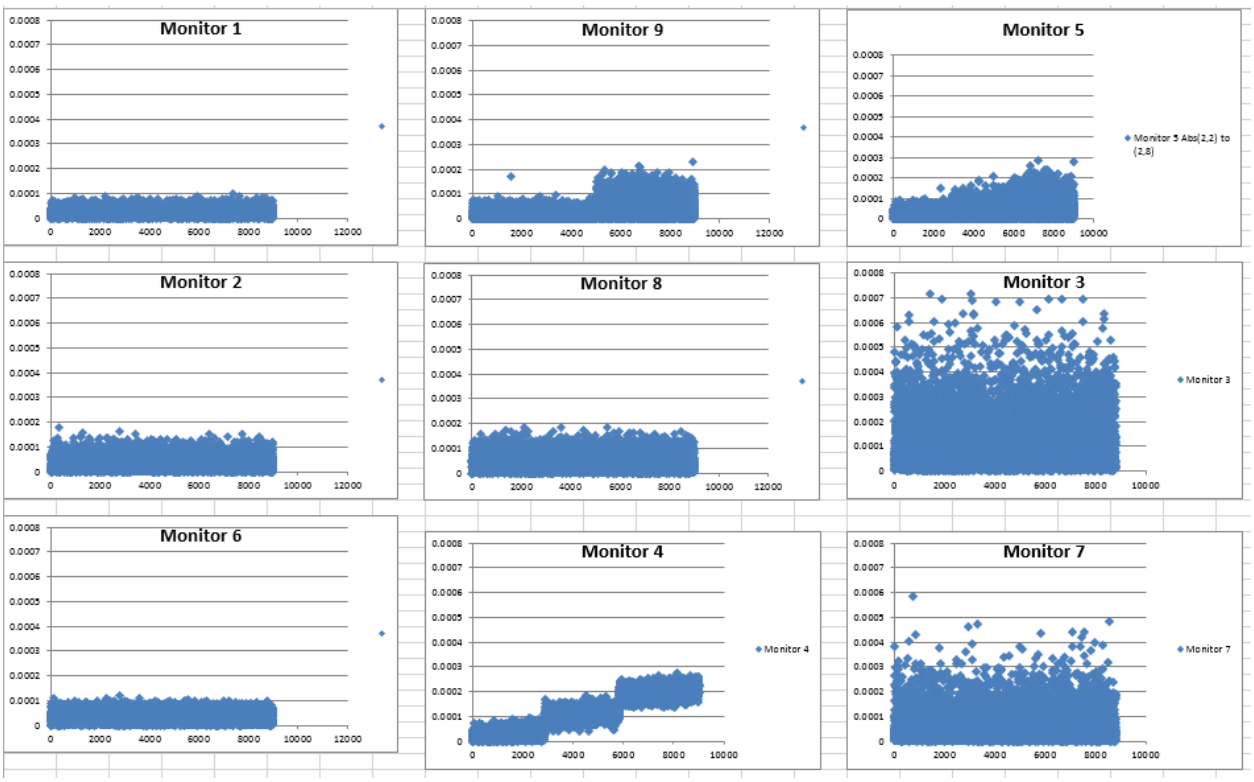

Figure 2

Figure 2 shows monitor readings. The monitors show behaviors of interest. Monitors 1, 2, 6, and 8 essentially show some Gaussian noise, with some variations in the noise and variations in the baselines. Monitor 4 displays a baseline that is shifted linearly over time, as we have gaps of three month between sample readings accounting for the stair-step effect. Monitor 5 gets noisier over time. Monitor 9 becomes noisier at the second month. Monitors 3 and 7 show more noise than the others.

2. *Now turn your attention to the chemicals themselves. Which chemicals are being detected by the sensor group? What patterns of chemical releases do you see, as being reported in the data? Limit your response to no more than 6 images and 500 words.*

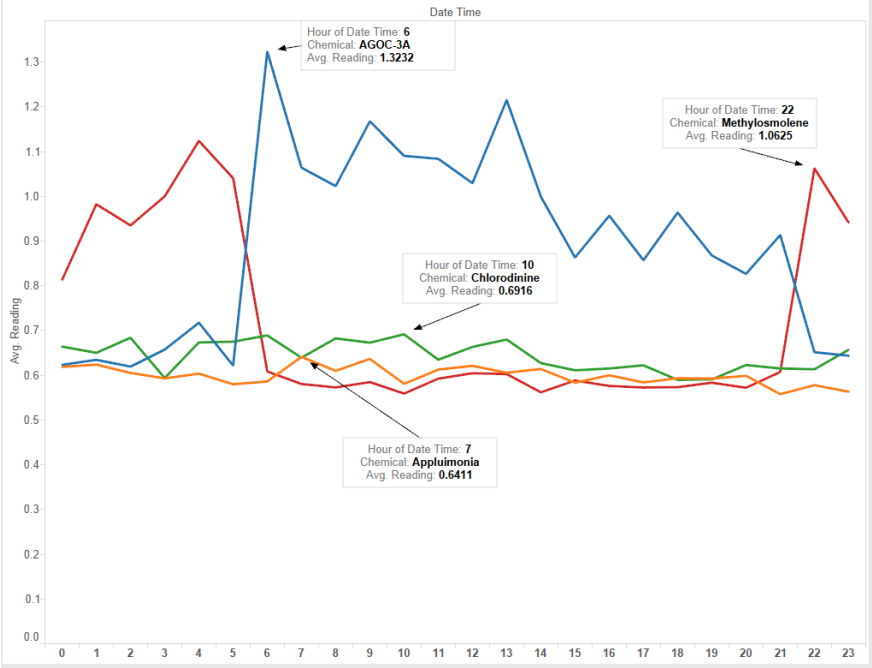

Figure 3

Figure 3 shows a simple line graph of releases over time, averaged on hours per day. We can see there are significant releases of AGOC-3A and Methylosmolene. Looking at the chemical description sheets, the Methylosmolene is a concern as the factories were supposed to have discontinued its use for environmentally safer chemicals. Additionally, the pattern for Methylosmolene release shows that it has very high releases during unusual hours (11PM-4AM).

Other ways of visualizing the chemical releases should be noted, as we want to encourage sound alternatives.

3. *Which factories are responsible for which chemical releases? Carefully describe how you determined this using all the data you have available. For the factories you identified, describe any observed patterns of operation revealed in the data. Limit your response to no more than 8 images and 1000 words.*

Factory 1 (Roadrunner) is the primary source for Chlorodinine emissions, but apparently at low levels.

Factory 2 (Kasios) emits Methylosmolene and AGOC-3A.

Factory 3 (Radiance) uses AGOC-3A extensively.

Factory 4 (Indigo) is the primary source for Appluimonia, but at low levels.

The task for the contestants is to clearly show this and how the emissions vary over time in their visualizations.

Additional Reviewer Considerations

MC2 Questions and Approach:

- Did the team specify all of the assumptions they made in addressing this problem? Were they reasonable?
- Did the visualizations play a central role in determining the solution to this problem?
- Did the submission present a primary suspect in emissions of the most hazardous chemical, Methylosmolene? If other chemicals were focal points, did they clearly explain why these were chosen over the more harmful ones?

MC2 Application of visual analytics:

- Did the team develop an innovative visual analytic tool? Alternatively, did they use an existing tool in an innovative way?
- Did visualizations enable the analysis process? Or did they merely illustrate conclusions? Did the submission rely more heavily on non-visual analytic approaches?
- Did their tool allow useful interactions?
- Did they use all the available data?
- Was the submission clear?