Data Visualization

Mystery at the Wildlife Preserve

Visual Analytics Science and Technology (VAST)
Mini Challenge 2

VAST Challenge

- An annual contest to advance the field of visual analytics
- The VAST Challenge is designed to help researchers understand how their software would be used in a novel analytic task and determine if their data transformations, visualizations, and interactions would be beneficial for particular analytic tasks.
- VAST Challenge problems provide researchers with realistic tasks and data sets for evaluating their software and advancing the field by solving more complex problems.

Visual Analytics Community
DISCOVERY THROUGH COLLABORATION AND SHARING

Challenge overview

- Mitch Vogel is a post-doc student studying ornithology and has been discovering that the number of the Rose-Crested Blue Pipit, a popular local bird is decreasing! The decrease is sufficiently significant that Mitch decided to undertake studies to identify the possible reasons. Mitch is gaining access to several datasets that may help him in his work.
- Mistford is a mid-size city is located to the southwest of a large nature preserve. The city has a small industrial area with four manufactories.

Mini Changeling 2

- The four factories in the industrial area are subjected to higher-than-usual environmental assessment, due to chemical releases. Gaseous effluent data from several sampling stations has been collected over three months, along with meteorological data (wind speed and direction), that could help Mitch understand what impact these factories may be having on the Rose-Crested Blue Pipit. These factories are supposed to be quite compliant with recent years' environmental regulations, but Mitch has his doubts that the actual data has been closely reviewed.

Goals/Questions

- Q1: 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?
- Q2: 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?
- Q3: 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.

Map

Generalgate 0
Entrance 0
Ranger Stop 0
Ranger Stop 1
Generalgate 1
Gate 0
Ranger Stop 2
Generalgate 2
Camping 0
Gate 1
Ranger Stop 3
Camping 1
Camping 8
Generalgate 0
Gate 2
Gate 3
Camping 2
Camping 3
Entrance 1
Entrance 2
Camping 4
Ranger Stop 4
Generalgate 4
Generalgate 5
Gate 4
Ranger Stop 5
Camping 5
Generalgate 6
Generalgate 7
Ranger Stop 6
Gate 5
Camping 7
Gate 6
4
5
9
3
Gate 7
2
Entrance 3
Roadrunner
Radiance
1
Kasio
6
Indigo
Camping 6
Gate 8
Entrance 4
8
7

Dataset

The following are the factory locations:

- Roadrunner Fitness Electronics: 89,27
- Kasios Office Furniture: 90,21
- Radiance ColourTek : 109,26
- Indigo Sol Boards: 120,22

The following are the sensor locations:

1:	62,21
2:	66,35
3:	76,41
4:	88,45
5:	103,43
6:	102,22
7:	89,3
8:	74,7
9:	119,42

Dataset

The Sensor data (provided in an Excel spreadsheet) contains 3 months of readings in the following format:

Chemical	Monitor	Date Time	Reading
Methylosmolene	3	4/1/16 0:00	2.68382
Methylosmolene	7	4/1/16 0:00	2.63064
Chlorodinine	3	4/1/16 0:00	1.25917
Chlorodinine	7	4/1/16 0:00	0.943983
AGOC-3A	3	4/1/16 0:00	0.722303

- Chemical: Appluimonia (Chemical App), Chlorodinine (Chemical C), Methylosmolene (Chemical Methy) and AGOC-3A (AGOC)
- Monitor: Which one of the nine sensors picking up the reading
- Reading: The air sensor detected amount in parts per million
- Date Time: The date and time of day of the reading, local time with no change for Daylight Savings.

Total: 79244 records

Dataset

The Meteorological data represents 3 months of readings in the following format:

Date	Wind Direction	Wind Speed (m/s)
4/1/16 0:00	190.5	4
4/1/16 3:00	203.3	5
4/1/16 6:00	201.1	5.2
4/1/16 9:00	204.9	4.1
4/1/16 12:00	207	3.6

Total: 708 records

Date: The date and time of the readings, local time with no change for Daylight Savings.
Wind Direction: The compass directions *where the wind is originating from*, using a north-referenced azimuth bearing where 360/000 is true north.
Wind Speed: The speed of the wind in meters per second.
Each of these reading is taken at the date and time provided.

Analysis Tool

- Python
- Tableau
- Excel

Q1: Sensor Performance

- Each sensor read 4 different chemical releases at every single hour, which should generate 216 readings (24 x 9) per day per chemical and 19872 readings in total for each chemical.

$$24 \times 9 \times (30+31+31) = 19872 \text{ readings}$$

Chemical	Date Time 2016
AGOC-3A	20,027
Appluimonia	19,810
Chlorodinine	19,809
Methylosmolene	19,597

Q1

- ***Using a python program, we found that a number of readings are completely missing at twelve or one a.m. for all sensors. This includes following 5 days: Apr. 2nd, 6th, Aug. 4th, 7th, and Dec. 2nd.***
- ***In addition, on Aug. 2nd and Dec. 7th, there are empty readings for most of sensors as Figure 1 shows.***

Sensors	Total Numbers	Time of missing readings			
		AGOC-3A	Appiulmonia,	Chlorodinine	Methylosmolene
1	28	4/2/16 0:00,	4/2/16 0:00,	4/2/16 0:00,	4/2/16 0:00,
		4/6/16 0:00,	4/6/16 0:00,	4/6/16 0:00,	4/6/16 0:00,
		8/2/16 0:00,	8/2/16 0:00,	8/2/16 0:00,	8/2/16 0:00
		8/4/16 0:00,	8/4/16 0:00,	8/4/16 0:00,	8/4/16 0:00,
		8/7/16 0:00	8/7/16 0:00	8/7/16 0:00	8/7/16 0:00
		12/2/16 0:00	12/2/16 0:00	12/2/16 0:00	12/2/16 0:00
		12/7/16 0:00	12/7/16 0:00	12/7/16 0:00	12/7/16 0:00
2	28	4/2/16 0:00,	4/2/16 0:00,	4/2/16 0:00,	4/2/16 0:00,
		4/6/16 0:00,	4/6/16 0:00,	4/6/16 0:00,	4/6/16 0:00,
		8/2/16 0:00,	8/2/16 0:00,	8/2/16 0:00,	8/2/16 0:00
		8/4/16 0:00,	8/4/16 0:00,	8/4/16 0:00,	8/4/16 0:00,
		8/7/16 0:00	8/7/16 0:00	8/7/16 0:00	8/7/16 0:00
		12/2/16 0:00	12/2/16 0:00	12/2/16 0:00	12/2/16 0:00
		12/7/16 0:00	12/7/16 0:00	12/7/16 0:00	12/7/16 0:00
3	26	4/2/16 0:00,	4/2/16 0:00,	4/2/16 0:00,	4/2/16 0:00,
		4/6/16 0:00,	4/6/16 0:00,	4/6/16 0:00,	4/6/16 0:00,
		8/2/16 0:00,	----------	----------	8/2/16 0:00
		8/4/16 0:00,	8/4/16 0:00,	8/4/16 0:00,	8/4/16 0:00,
		8/7/16 0:00	8/7/16 0:00	8/7/16 0:00	8/7/16 0:00
		12/2/16 0:00	12/2/16 0:00	12/2/16 0:00	12/2/16 0:00
		12/7/16 0:00	12/7/16 0:00	12/7/16 0:00	12/7/16 0:00
4	28	4/2/16 0:00,	4/2/16 0:00,	4/2/16 0:00,	4/2/16 0:00,
		4/6/16 0:00,	4/6/16 0:00,	4/6/16 0:00,	4/6/16 0:00,
		8/2/16 0:00,	8/2/16 0:00,	8/2/16 0:00,	8/2/16 0:00
		8/4/16 0:00,	8/4/16 0:00,	8/4/16 0:00,	8/4/16 0:00,
		8/7/16 0:00	8/7/16 0:00	8/7/16 0:00	8/7/16 0:00
		12/2/16 0:00	12/2/16 0:00	12/2/16 0:00	12/2/16 0:00
		12/7/16 0:00	12/7/16 0:00	12/7/16 0:00	12/7/16 0:00
5	28	4/2/16 0:00,	4/2/16 0:00,	4/2/16 0:00,	4/2/16 0:00,
		4/6/16 0:00,	4/6/16 0:00,	4/6/16 0:00,	4/6/16 0:00,
		8/2/16 0:00,	8/2/16 0:00,	8/2/16 0:00,	8/2/16 0:00
		8/4/16 0:00,	8/4/16 0:00,	8/4/16 0:00,	8/4/16 0:00,
		8/7/16 0:00	8/7/16 0:00	8/7/16 0:00	8/7/16 0:00
		12/2/16 0:00	12/2/16 0:00	12/2/16 0:00	12/2/16 0:00
		12/7/16 0:00	12/7/16 0:00	12/7/16 0:00	12/7/16 0:00
6	27	4/2/16 0:00,	4/2/16 0:00,	4/2/16 0:00,	4/2/16 0:00,
		4/6/16 0:00,	4/6/16 0:00,	4/6/16 0:00,	4/6/16 0:00,
		--------	8/2/16 0:00,	8/2/16 0:00,	8/2/16 0:00
		8/4/16 0:00,	8/4/16 0:00,	8/4/16 0:00,	8/4/16 0:00,
		8/7/16 0:00	8/7/16 0:00	8/7/16 0:00	8/7/16 0:00
		12/2/16 0:00	12/2/16 0:00	12/2/16 0:00	12/2/16 0:00
		12/7/16 0:00	12/7/16 0:00	12/7/16 0:00	12/7/16 0:00
7	26	4/2/16 0:00,	4/2/16 0:00,	4/2/16 0:00,	4/2/16 0:00,
		4/6/16 0:00,	4/6/16 0:00,	4/6/16 0:00,	4/6/16 0:00,
		--------	--------	8/2/16 0:00,	8/2/16 0:00
		8/4/16 0:00,	8/4/16 0:00,	8/4/16 0:00,	8/4/16 0:00,
		8/7/16 0:00	8/7/16 0:00	8/7/16 0:00	8/7/16 0:00
		12/2/16 0:00	12/2/16 0:00	12/2/16 0:00	12/2/16 0:00
		12/7/16 0:00	12/7/16 0:00	12/7/16 0:00	12/7/16 0:00
8	26	4/2/16 0:00,	4/2/16 0:00,	4/2/16 0:00,	4/2/16 0:00,
		4/6/16 0:00,	4/6/16 0:00,	4/6/16 0:00,	4/6/16 0:00,
		--------	8/2/16 0:00,	8/2/16 0:00,	--------
		8/4/16 0:00,	8/4/16 0:00,	8/4/16 0:00,	8/4/16 0:00,
		8/7/16 0:00	8/7/16 0:00	8/7/16 0:00	8/7/16 0:00
		12/2/16 0:00	12/2/16 0:00	12/2/16 0:00	12/2/16 0:00
		12/7/16 0:00	12/7/16 0:00	12/7/16 0:00	12/7/16 0:00
9	28	4/2/16 0:00,	4/2/16 0:00,	4/2/16 0:00,	4/2/16 0:00,
		4/6/16 0:00,	4/6/16 0:00,	4/6/16 0:00,	4/6/16 0:00,
		8/2/16 0:00,	8/2/16 0:00,	8/2/16 0:00,	8/2/16 0:00
		8/4/16 0:00,	8/4/16 0:00,	8/4/16 0:00,	8/4/16 0:00,
		8/7/16 0:00	8/7/16 0:00	8/7/16 0:00	8/7/16 0:00
		12/2/16 0:00	12/2/16 0:00	12/2/16 0:00	12/2/16 0:00
		12/7/16 0:00	12/7/16 0:00	12/7/16 0:00	12/7/16 0:00

Q1

Monitor status
pt.1
Hour of Date Time
Chemical
Monitor
AGOC-3A
Appluimonia
Chlorodinine
Methylosm...

Q1

Monthly Heapmap

Q1

Even when there is no wind blowing to the sensors according to figure 3 (where gray circle indicates wind direction in y axis along with time in hour in x axis), there is still a consistent level of release detected by sensor 3&4. We conclude it should be the chemicals from the environment

Q2: Chemical Release Pattern

Q2

Chemical pattern
Date Time
Week of Da..
0
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
March 27, 2016
April 3, 2016
April 10, 2016
April 17, 2016
April 24, 2016
July 31, 2016
August 7, 2016
August 14, 2016
August 21, 2016
August 28, 2016
November 27, 2016
December 4, 2016
December 11, 2016
December 18, 2016
December 25, 2016
Chemical
AGOC-3A
Applumonia
Diphosphine
Methylparadiene

Q2

Monitor status
Hour of Date Time
Chemical
Monitor
ASOC-3A
Methylosmolene

Q3: Factory Responsibility

$\cos\Delta = V \cdot W$

			Senor 1				Sensor 2				Sensor 3			
Date	Wind Direction	Wind Speed (m/s)	Roadrunner	Kasios	Radiance	Indigo	Roadrunner	Kasios	Radiance	Indigo	Roadrunner	Kasios	Radiance	Indigo
4/1/16 0:00	190.5	4	0.00	0.00	0.00	0.00	0.15	0.34	0.02	0.05	0.60	0.34	0.02	0.05
4/1/16 3:00	203.3	5	0.00	0.00	0.00	0.00	0.00	0.12	0.00	0.00	0.40	0.12	0.00	0.00
4/1/16 6:00	201.1	5.2	0.00	0.00	0.00	0.00	0.00	0.16	0.00	0.00	0.44	0.16	0.00	0.00
4/1/16 9:00	204.9	4.1	0.00	0.00	0.00	0.00	0.00	0.09	0.00	0.00	0.38	0.09	0.00	0.00

Peak Readings

For every 100 readings,

we calculated an average using Python;

If a reading is above the average,

it is a peak reading;

else

it is ignored.

Q3

Factory release
Monitor
Day of Date
Possibility
Reading
Monitor
(All)
1
2
3
4
5
6
7
8
9
Chemical
(All)
AGOC-3A
Appluimonia
Chlorodinine
Methylosmolene
Factory
Indigo
Kasios
Radiance
Roadrunner
ATTR(Possibility)
0.4909
0.6000
0.7000
0.8000
0.9000
1.0000
Chemical
AGOC-3A
Appluimonia
Chlorodinine
Methylosmolene

Q3

Factory release
Monitor
Hour of Date
Possibility
Reading
3
5
6
9
1.0
0.8
0.6
9
8
7
6
5
A...
A...
A...
Augus...
A...
Dece...
Dece...
Dece...
Dece...
Dece...
Monitor
(All)
1
2
3
Chemical
(All)
AGOC-3A
Appluimonia
Chlorodinine
Factory
Indigo
Kasios
Radiance
Roadrunner
ATTR(Possibility)
0.4924
0.6000
0.7000
0.8000
Chemical
Appluimonia

Q3

Indigo mostly likely is responsible for Appluimonia's release; Radiance mainly released AGOC-3A; Kasios and Roadrunner both released all the other chemicals.

	AGOC-3A	***Appluimonia***	***Chlorodinine***	***Methylosmolene***
Indigo				
Kasios				
Radiance				
Roadrunner				

Factory Indigo Release Pattern

Possibility for each Peak Time Weekday broken down by Date Year. Color shows details about Chemical. Size shows sum of Reading. Details are shown for Factory. The data is filtered on Reading and Peak Time Hour. The Reading filter ranges from 5 to 80.4121. The Peak Time Hour filter keeps 24 of 24 members. The view is filtered on Chemical, Possibility and Factory. The Chemical filter keeps AGOC-3A, Appluimonia, Chlorodinine and Methylosmolene. The Possibility filter ranges from 0.4900 to 1.0000. The Factory filter keeps Indigo.

Factory Radiance Release Pattern

Possibility for each Peak Time Weekday broken down by Date Year. Color shows details about Chemical. Size shows sum of Reading. Details are shown for Factory. The data is filtered on Reading and Peak Time Hour. The Reading filter ranges from 5 to 80.4121. The Peak Time Hour filter keeps 24 of 24 members. The view is filtered on Chemical, Possibility and Factory. The Chemical filter keeps AGOC-3A, Appluimonia, Chlorodinine and Methylosmolene. The Possibility filter ranges from 0.4900 to 1.0000. The Factory filter keeps Radiance.

Factory Kasios Release Pattern

Possibility for each Peak Time Weekday broken down by Date Year. Color shows details about Chemical. Size shows sum of Reading. Details are shown for Factory. The data is filtered on Reading and Peak Time Hour. The Reading filter ranges from 5 to 80.4121. The Peak Time Hour filter keeps 24 of 24 members. The view is filtered on Chemical, Possibility and Factory. The Chemical filter keeps AGOC-3A, Appluimonia, Chlorodinine and Methylosmolene. The Possibility filter ranges from 0.4900 to 1.0000. The Factory filter keeps Kasios.

Factory Roadrunner Release Pattern

Possibility for each Peak Time Weekday broken down by Date Year. Color shows details about Chemical. Size shows sum of Reading. Details are shown for Factory. The data is filtered on Reading and Peak Time Hour. The Reading filter ranges from 5 to 80.4121. The Peak Time Hour filter keeps 24 of 24 members. The view is filtered on Chemical, Possibility and Factory. The Chemical filter keeps AGOC-3A, Appluimonia, Chlorodinine and Methylosmolene. The Possibility filter ranges from 0.4900 to 1.0000. The Factory filter keeps Roadrunner.

New Data Visualization

Future Plan

- Large Data scale (more than 79244 records)
- Large number of dimensions: Chemical, Factory, Sensor, Time (per hour, per day, per week, per month)
- Aggregation methods for data compression
- Exploiting high-dimensional data display techniques in VR.

Publication plan

- IEEE VIS two pages paper;
- IEEE Big Data in Dec.

Acknowledgement

Ronald Ashley

Rummel Jessamy

Acknowledgement

- BNL collaborator Wei Xu
- OEP, Noel and Sal

Acknowledgement

- My student team

- Questions???

Paper published with student

- Bo Sun, **Daniel Vozanilek**. Tumor Vascular Permeability Measurement based on Color Image Analysis. *Proceedings of 2012 International Conference on Systems and Informatics (ICSAI 2012)*, Yantai, China, May 19th -21st, 2012 (2003-2007): IEEE Xplore.

Daniel Vozenilek

Web/Java Developer at Vanguard

Malvern, Pennsylvania | Computer Software

Previous Lincoln University, The Vanguard Group, Visual Sound

Education Lincoln University

Send Daniel InMail

conn

Paper published with student

- Bo Sun. "An image processing-based method for quantification of microvasculature". *The 2014 IEEE international conference on multimedia and expo.*, Chengdu, china, July 2014.