VAST Challenge 2014

Mini-Challenge 3 Guide for Reviewers

Mini-Challenge 3 provides streaming data for participants to analyze. This data includes both microblog data and emergency call center data.

# The Task

The contestants are asked to provide answers to three questions, using visual analytic techniques to support their investigation:

**MC3.1** – Within 3 hours after start the Segment 3 data stream, send an email to [VASTChal2014MC3@vacommunity.org](mailto:VASTChal2014MC3@vacommunity.org) containing

1. An image showing the streaming data in your visual analytics tool. In this image, identify an event of interest that you intend to investigate further.
2. the content of the final message in the data stream

**MC3.2** – Describe the timeline of up to five major events that you discover in the streaming data. This timeline should include information from all three segments of the data stream if needed. Use specific microblog records and call center data to support your description, but do not simply mimic back the data stream. Provide a concise description of important participants, locations and durations. Focus your response on the events themselves, rather than on the individuals reporting the events. Please limit your answer to no more than ten images and 1500 words.  
  
**MC3.3** – Select one of your five major events from question MC 3.2 that you consider to be most likely to provide additional clues to the investigation of the GASTech disappearances. Describe the roles of the participants. Describe how other events you identified in MC3.2 may have influenced your selected event. Provide a hypothesis and evidence as to whom you suspect as being directly involved in the GASTech disappearances, either as perpetrators or victims. Please limit your response to no more than five images and 500 words.

# The Data

The data for MC3 consists of 3 time-windowed segments of data provided in a streaming context. The data covers a bit more than 4.5 hours on the evening of January 23. The data is a collection of micro-blog and call center records that occur over time. Segment 1 was provided both as a stream and as an initial download of csv records. Segments 2 and 3 were provided only as streaming data. Segments 1 and 2 could be replayed as many times as necessary. Segment 3 could only be played once.

# Answers to Questions

## Question MC3.1

*Within 3 hours after start the Segment 3 data stream, send an email to* [*VASTChal2014MC3@vacommunity.org*](mailto:VASTChal2014MC3@vacommunity.org) *containing*

1. *An image showing the streaming data in your visual analytics tool. In this image, identify an event of interest that you intend to investigate further.*
2. *the content of the final message in the data stream*

Participants submitted their answers were instructed to include their answers in to this question in an email to the above email address and to include their answers in this form as well. The answer should correctly indicate one of the final messages in the stream, as shown in the table below. The last 3 messages are shown in the table. Extra messages are provided for clarity as it is suspected that at least one submission contained a message that was not the last message.

|  |  |  |
| --- | --- | --- |
| DateTime-YYYYMMDDHHMISS | Author | Message Content |
| 20140123213400 | lindyT | RT @KronosStar There has been an explosion from inside the apartment building. Several people are down. #KronosStar #DancingDolphinFire #AFDHeroes |
| 20140123213400 | dolls4sale | RT @redisrad What was that? #boom |
| 20140123213445 | worldWatcher | RT @CentralBulletin explosion heard at dancing dolphin apartment #Abila #CentralBulletin |

The 2 top messages have the same timestamp 1/23/2014 21:34:00 – so may be sent with either message first. The last message in the table is the actual final message at 1/23/2014 21:34:45.

In addition, the answer should indicate that they have identified a significant event in the data, and the image included should demonstrate that they used their visual analytics tool to analyze the streaming data, rather than capturing all data to analyze in a batch.

## Question MC3.2

*Describe the timeline of up to five major events that you discover in the streaming data. This timeline should include information from all three segments of the data stream if needed. Use specific microblog records and call center data to support your description, but do not simply mimic back the data stream. Provide a concise description of important participants, locations and durations. Focus your response on the events themselves, rather than on the individuals reporting the events. Please limit your answer to no more than ten images and 1500 words.*

Reviewers, please note that in the microblog messages, the message will occur seconds to minutes after the event itself. They may also be re-transmitted multiple times. News sources have a higher proportion of retransmission than do other microblog sources. Re-transmitted microblog messages begin with “RT”.

The following are the major events present in the data.

1. **Fire at the Dancing Dolphin apartments** – A 6 story apartment building on the east side of town catches on fire. The fire is called to emergency services at 18:40. The fire continues through the entire duration of the data streams, culminating with an explosion at 21:30 after it seemed like the fire was under control.
2. **POK rally in the park** – The Protectors of Kronos are holding a rally in a park on the west side of town – near the government buildings. It begins at 17:00. Various speakers and songs are played until 19:05. Viktor-E continues to play music in the park as a free concert after the formal rally is complete.
3. **Hit and Run** – A black van clips a car on Souliou. It then hits a cyclist on the 500 block of Schaber between Eira and Edessis. The cyclist is okay – helped out by people from Brew’ve Been Served and a witness.

A witness to the car that was hit is located at 19:47. The witness knows where the black van came from.

The van continues away from the accident – a partial plate is known “##L829”

This event continues into the stand-off – so may be reported as 1 contiguous event.

1. **Stand-off** - The van runs a red light at 19:35. A police car sees this and combined with the partial license plate pursues the van.

The van enters a parking lot next to GelatoGalore – and start a confrontation that becomes a standoff at 19:40.

An occupant of the van claims to have hostages. Much talking and arrival of SWAT, eventually a different occupant of the van comes out and gives up – ending the standoff peacefully at 21:13.

Note : The streaming data does not actually reveal the names of the drivers or hostages.

Isia Vann and Inga Ferro were the van drivers.

Hostages in the van are Rachel Pantanal and Ada Campo-Corrente.

1. **Safe house found** – The location of the black van is found and searched; finding much interesting evidence that reveals it was a kidnapper’s safe house at 21:22 – In the call center data.

## Question MC3.3

*Select one of your five major events from question MC 3.2 that you consider to be most likely to provide additional clues to the investigation of the GASTech disappearances. Describe the roles of the participants. Describe how other events you identified in MC3.2 may have influenced your selected event. Provide a hypothesis and evidence as to whom you suspect as being directly involved in the GASTech disappearances, either as perpetrators or victims. Please limit your response to no more than five images and 500 words.*

The most likely event to be reported by contestants is the standoff, possibly combined with the hit and run.

The drivers of the black van were in a location near the Dancing Dolphin Fire. With the impending evacuation, they panicked and left the location in a rush. They side-swiped a car. They then proceed down the street and hit a bicyclist right outside Gelato Galore – a partial license plate is reported.

As they run from that scene, they run a light. They are chased by police and it turns into a stand-off.

The drivers claim to have hostages. It is a likely conclusion that the hostages may be part of the missing GASTech employees. Some microblog messages speculate that the hostages are the missing GASTech people.

Extrapolating that the drivers are part of the GASTech kidnapping, the source of their flight is likely to provide insight into the kidnapping.

The streaming data does not actually reveal the names of the drivers or hostages.

When the police find the source location of the black van, they call for crime scene investigation then they later call for additional support at the crime scene. It could be extrapolated that the source location of the black van was a kidnapping safe house.

Other possible answers

* The POK rally was staged as a distraction for the kidnappers – so potentially there would be POK members there that could be detained and questioned. There is no specific evidence pointing to this – though there are microblog messages with a definite negative tone towards the POK.
* The Dancing Dolphin fire was set on purpose and was intended to destroy evidence. The cause of the fire is not discussed in the data.

Answers that have a logical progression and some data to support that chain of reasoning should be accepted. It should be noted that the microblog data is a mix of news outlets, official accounts from the Abila Fire and Police, and individuals. As with most social media data, not all messages are true or accurate. It is possible that the contestants accounted for this and have placed more analytic weight on the ‘official’ sources or down played single data points that have no corroboration.

Contestants that have done other the mini-challenges may reference evidence from the other challenges. Patterns on life from MC2 may yield additional insight into the fact that there is a location near the Dancing Dolphin fire that is of interest. Some person-of-interest names would only be known with information from other mini-challenges.