

VAST Challenge 2015: Reviewer Guide

This document provides information to support peer review of submissions to VAST Challenge 2015. This document includes sections for each of the two mini-challenges and the Grand Challenge. It 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+2015>.

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 vast-challenge@ieee.org for assistance.

Challenge Summary

Please note: All organizations, locations, and events are fictitious.

Setting

The Challenge is set in “DinoFun World”, a modest-sized amusement park, sitting on about 215 hectares and hosting thousands of visitors each day. It has a small town feel, but it is known for its exciting rides and events. In June 2014, the park hosted a weekend tribute to Scott Jones, internationally renowned football (“soccer,” in US terminology) star. Scott Jones is from a town near DinoFun World. He was a hometown hero, with thousands of fans who cheered his success as if he were a family member. To celebrate his years of stardom in international play, DinoFun World declared “Scott Jones Weekend” on Friday, June 6 through Sunday, June 8, 2014. Scott was scheduled to appear in two stage shows each day to talk about his life and career. In addition, a show of memorabilia related to his career was displayed in the park’s Pavilion.

Data Provided

The primary source of data for the participants is the DinoFun World app. All visitors to the park (except very young children) must use the app to check in to the park and its rides and to communicate with fellow visitors. If visitors do not have compatible phones, the park loans devices to them. Visitors are assigned IDs that remain active for the length of their ticket, which could be multiple days.

The park is equipped with sensor beacons that record movements within the park. Sensors are sensitive within a 5m x 5m grid cell. All pathways in the park are covered by these sensors, as are the ride check-in locations. Locations are not recorded while people are on rides or inside attractions (including restaurants, stores, and restrooms).

App users may send text messages to anyone within their own designated group (for example, a family could have its own group). An app user may also make a friend at the park where they can send and receive texts.

In addition to the movement data (Mini-Challenge 1) and communication data (Mini-Challenge 2) from the app, participants in the VAST Challenge were provided a detailed map of the park, a description of the data, a DinoFun World web site excerpt, and a news article about an event that occurred during the weekend.

Challenge Goals

Participants were tasked with using visual analytics to understand how people moved and communicated throughout DinoFun World and to note patterns and their changes over time. Understanding the pattern changes helps to inform the teams about deviations and other activities that are relevant to the crimes. In addition, it allows teams to identify issues with park operations.

Ground Truth

The following description was not provided to participants. This describes the main event reflected in the data set. On Sunday, June 8, between 9:30 and 11:32 a.m., the display of Scott Jones's soccer memorabilia in the Creighton Pavilion at DinoFun World was vandalized. The display was open to the ticketed visitors of the park from the hours of 8:00 through 9:30 a.m., and then the Pavilion was closed to free the park security guards for crowd control duty at one of Scott's shows. The security guards returned to the Pavilion and re-opened it at approximately 11:30 a.m. Police arrived at the park at 12:04 p.m. and observed that a fire extinguisher had been used to break open several glass display cases. A number of the items of memorabilia had been damaged or had been defaced with black spray paint. Spray paint had been used to write derogatory statements about Scott Jones on the display and throughout the Pavilion. There were no signs of forced entry into the Pavilion. Police were provided a list of the items in the display and notably missing from the items on display was Scott Jones's Olympic medal.

The criminal acts are performed by Eddie Smith and two anonymous henchmen. Eddie was a teammate of Scott Jones in high school, but was left behind as Scott became increasingly celebrated for his outstanding skills. Eddie's jealousy finally overcame him, which led to the plot and Sunday's events. Details of Eddie's plot are described as part of the answers below.

It is not expected that participants identify Eddie by name, but we hope they will recognize him by his ID number in the data: 1983765.

Challenge Questions and Potential Answers

The sections below outline the questions and potential answers for the VAST Challenge. Please refer to the section corresponding to the mini-challenge for which you are providing reviews.

Please note that reviewers are not expected to review for accuracy of answers using this guide. The answers provided below represent the patterns that were embedded in the data. There are almost

certainly additional patterns detectable in the data that were not embedded intentionally. If participants provide a detailed, data-driven rationale for their answers, they should be considered equally correct as answers provided below. If you have questions about whether the data support a particular assertion, please contact vast-challenge@ieeevis.org for assistance.

Note:

- This year, participants are permitted (but not required) to use all of the data for both mini-challenges to solve a given mini-challenge. Participants were asked to state which data sets they used to determine their answers. Answers that are informed by both the movement and the communications data will likely have more complexity and detail than those informed by only one data source. However, participants should not be penalized if they chose to use only one data source.
- In addition, it was discovered near the end of the challenge period that a portion of the movement data for Friday, June 6 from around 8 p.m. to park closing was missing from the initial data release. Participants had the option to include this data in their analysis or to exclude it.

Mini-Challenge 1: Park Movement

Participants are given movement tracking information for all of the paying park visitors for June 6-8, 2014.

MC 1.1. Characterize the attendance at the park on this weekend. Describe up to twelve different types of groups at the park on this weekend.

- a. How big is the group type?**
- b. Where does this type of group like to go in the park?**
- c. How common is this type of group?**
- d. What are your other observations about this type of group?**
- e. What can you infer about the group?**
- f. If you were to make one improvement to the park to better meet this group's needs, what would it be?**

Please limit your response to no more than 12 images and 1000 words.

Participants may use multiple techniques to characterize groups. We anticipate that most groups will be characterized based on size and favored activities. Questions a. through c. are relatively straightforward and data driven. For question d., we are hoping to see responses that notice particular data patterns such as staying together or splitting into subgroups during their visit. For question e., we hope that participants take a guess at what kind of group it is, such as a family with small children or a high school group. For question f., any thoughtful data-driven response is appropriate.

During the data generation process, approximately 27 low-level groups types were defined. Each of those group types are present in the park certain level per day. Key variables include

- The number of instances of a particular group in a day
- The composition of the group (adults, teens, young children , etc.)
- Most likely entrance and exit
- Hours of park attendance, and frequency of departure and return during the day
- Number of days they come to the park

Individual groups also have chances to engage in activities such as attending the Scott Jones show, shopping, searching the park specific items they wish to purchase. Participants may identify the some of the items mentioned in MC1.3 here. However, we expect that the majority of the answers will characterize the overall group behaviors, which are summarized in the following table.

Group Type	Size	Park Areas Visited	Frequency of Occurrence	Other Observations	Potential Inference
Couples	2 people	All places, small chance of kiddie rides	Uncommon	Stays full day, may return multiple days	Two adults touring together
Groups with Subgroups	30-80 people. Arrive as a large group and split into subgroups of 1-2 adults, 5-9 kids	All places, small chance of kiddie rides	Common	Arrive together and breaks into smaller groups. Small number of these types of groups return multiple days	Maybe a school or touring group
Adult Groups	1-4 people	All places, no kiddie rides	Uncommon	Arrive and tour as a group. Visits beer garden	Adults touring park together
Family Groups	1-2 adults, 1-4 kids	General rides and kiddie rides	Common	Arrive as a single group and stay together	A Family-type group
Small Group that splits	1-3 adults, 1-5 teens	All rides	Moderately common	Arrive as a single group and split apart	Family-type group with older kids who wander independently
Slow Tourists	15 adults plus 15 helpers	Few rides, watch Scott show	Rare: 1 group of 30 on Friday	Movement rate is remarkably slow	Wheelchair-bound group from an assisted living facility
Tour Group	30-44 adults	Most rides, unlikely kiddie rides	Moderately common	Arrive as a single group and stay together	Large touring group
Singletons	1 adult	No rides or shows	8 on Friday, 10 on Saturday, 13 on Sunday	Do not ride, but pause for long periods	Photographers or performers

MC 1.2. Are there notable differences in the patterns of activity on in the park across the three days? Please describe the notable difference you see. Please limit your response to no more than 3 images and 300 words.

Participants may identify several changes in patterns of activity. The following are a few of the patterns that may be identified:

- Friday's volume of attendees is nearly half of Saturday or Sunday. Attendance increases markedly throughout the weekend.
- Several groups try to attend the second Scott Jones on Sunday only to discover that it has been canceled.
- More people leave during the day on Sunday than Friday/Saturday. This is a subtle signal, as it has some relation to the number of people present in the park each day.
- Other notable differences in movement patterns may be related to park operational events related in the solution for MC 1.3.

MC 1.3. What anomalies or unusual patterns do you see? Describe no more than 10 anomalies, and prioritize those unusual patterns that you think are most likely to be relevant to the crime. Please limit your response to no more than 10 images and 500 words.

The following are instances of unusual movement-pattern related instances that may be reported:

1. Closed rides: park visitors will not queue for these rides
 - The Red Train "Scholtz Express" (Ride 20) closed Saturday 13:23 - Saturday 13:50
 - The Flying Tyrandrienkos (Ride 12) closed Sunday 10:13 - Sunday 11:00
2. Galactosaurus Rage (Ride 2) has operational problems. Friday 8:00 Ride open but not operating; it resumes operation at 9:00
 - Friday 19:02 Ride stuck, people on ride do not get off
 - Friday 19:31 Ride unstuck, people get off; ride closed
 - Friday 20:01 Normal operation resumes
3. Food poisoning from Granite Slab Pizza (Food Stand Attraction 58)
 - Friday 12:01 Problems begin. Some people become ill after eating at the stand. There is a chance that people go to the first aid station after eating there.
 - Saturday 8:00 Open, normal operation
4. Scott Jones show that occurs at the Grinosaurus Stage (Attraction 63). Park visitors move into this attraction prior to show time and depart when the show is finished. The area is open at the following times:
 - Friday 8:00 - Friday 11:00
 - Friday 13:00 - Friday 16:00
 - Saturday 8:00 - Saturday 11:00

- Saturday 13:00 - Saturday 16:00
 - Sunday 8:00 - Sunday 11:00
5. The Creighton Pavilion (Attraction 32) opens and closes regularly. The Pavilion closes before each of the Scott Jones shows to allow security to move to the show grounds. Consequently, the Pavilion is closed at the following times.
 - Friday 9:30 – 11:30
 - Friday 14:30 – 16:30
 - Saturday 9:30 – 11:30
 - Saturday 14:30 – 16:30 Sunday 9:30 -11:30
 - Sunday 12:00 closed for the remainder of the day due to the vandalism.
 6. It may be noted that sometimes a single visitor goes to rides but never checks in. This may be innocuous. One example may be a photographer at the park to take pictures.
 7. Visitor 1187304 represents a “marathon roller coaster rider.” This person rides roller coaster 1, the Wrightiraptor Mountain repeatedly without checking in again. He rides 10 times, goes to the restroom, and rides seven more times. After that, he goes to location 62, the Liggament Fix-Me-Up (first aid station) and exits the park.
 8. Visitors 258464 and 74766 dare each other to ride one of the scarier rides. They return to the ride frequently on Saturday morning but do not actually work up the courage to ride it until 11:30.
 9. Visitors 657863, 1412235, 103006, 1937834, and 313073 exhibit unusual behaviors. They essentially spend Friday in the beer garden. Then they ride the train and all but 657863 leave the park. Visitor 657863 spends the night in the park and does not leave until Saturday morning.
 10. Detailed movements for Eddie Smith (1983765) and his accomplices (1089132 and 1723967) may be identified. The perpetrators perform surveillance on Friday and Saturday, and Eddie performs the vandalism while his accomplices stand watch. See the GC answers for additional information on these people.
 11. Starting at 20:18 on Saturday, visitor 1983765 (Eddie Smith) has duplicate entries of movement placing him at a different parts of the park. This is not repeated on Sunday. See the GC answers for more information about this anomaly.
 12. Detailed movements of Scott Jones and his followers may be reported. Scott’s movement is not tracked, but members of his entourage (bodyguard, assistants, and some close friends) travel through the park on the same path each day with Scott on the way to and from his show. Scott’s entourage is 1080969, 1600469, 1629516, 1781070, 1935406, 521750, and 644885.

Mini-Challenge 2: Park Communications

Participating teams were given “in-app” communication data for June 6-8, 2014. This includes communications among the paying park visitors, as well as communications between the visitors and park services. In addition, the data also contained records indicating when the user sent a text to an external party.

MC 2.1. Identify those IDs that stand out for their large volumes of communication. For each of these IDs:

- a. Characterize the communication patterns you see.**
- b. Based on these patterns, what do you hypothesize about these IDs?**

Please limit your response to no more than 4 images and 300 words.

The two IDs that stand out for large numbers of communications are 1278894 and 839736.

ID 1278894 sends and receives messages to multiple park visitors. This ID send messages, and others respond, at regular intervals for a period, and then the pattern ceases for an interval. Not all IDs in the park receive communications from this ID. The 1278894 pattern could be a status check of the device/app, an advertisement beacon, or an intermittent game. Based on the DinoFun World web site provided, participants could correctly hypothesize that this is a park-wide trivia game.

ID 839736 communicates with messages to several different park visitors throughout the day at no fixed interval. The 839736 pattern could also be a status check of the device, an advertisement beacon, or some form of communication with the park management. The ground truth for this pattern is communication with the park’s help desk.

MC 2.2. Describe up to 10 communications patterns in the data. Characterize who is communicating, with whom, when and where. If you have more than 10 patterns to report, please prioritize those patterns that are most likely to relate to the crime. Please limit your response to no more than 10 images and 1000 words.

Participants may identify multiple patterns, such as the following:

1. Group leaders send bulk messages to their groups. For example, visitor146240 sent messages to a large group at 13:34:26 on Friday. Some of their message recipients respond. The pattern is repeated at other times during the day.
2. When people are standing in a ride queue or at a food and drink attraction, they have an increased possibility of making friends with people around them who did not accompany them to the park. They can communicate with these friends during the rest of their time in the park.
3. There is an increased likelihood that people will send external messages when Scott Jones is in the park (8:45-11:35 each day and 13:45-16:30 on Friday and Saturday), specifically when they are near Scott as he travels through the park. (See the map in GC1 below for Scott’s daily path.)
4. There is reduced external messaging when Scott Jones shows are in progress.
5. There is an increase in messages to the traveler’s group, the help desk, and external contacts when the Scott Jones memorabilia vandalism is discovered. This occurs on Sunday from 11:30-12:00.
6. There is an increase in external messages when police arrive at the park just after noon on Sunday, specifically for people near the police as they travel through the park.

MC 2.3. From this data, can you hypothesize when the vandalism was discovered? Describe your rationale. Please limit your response to no more than 3 images and 300 words.

The crime occurs during the first (and only) show on Sunday. People visiting the Pavilion shortly after the show increase their contact with external contacts, other group members, and the help desk, supporting the hypothesis that the vandalism was discovered immediately after the first show. Additional communication, particularly external communication, occurs as the police moves through the park to investigate shortly after 12 noon.

Grand Challenge: Uncovering a Nefarious Plot

Information pulled from both movement and communications data may be combined to assist in describing the weekend's activities.

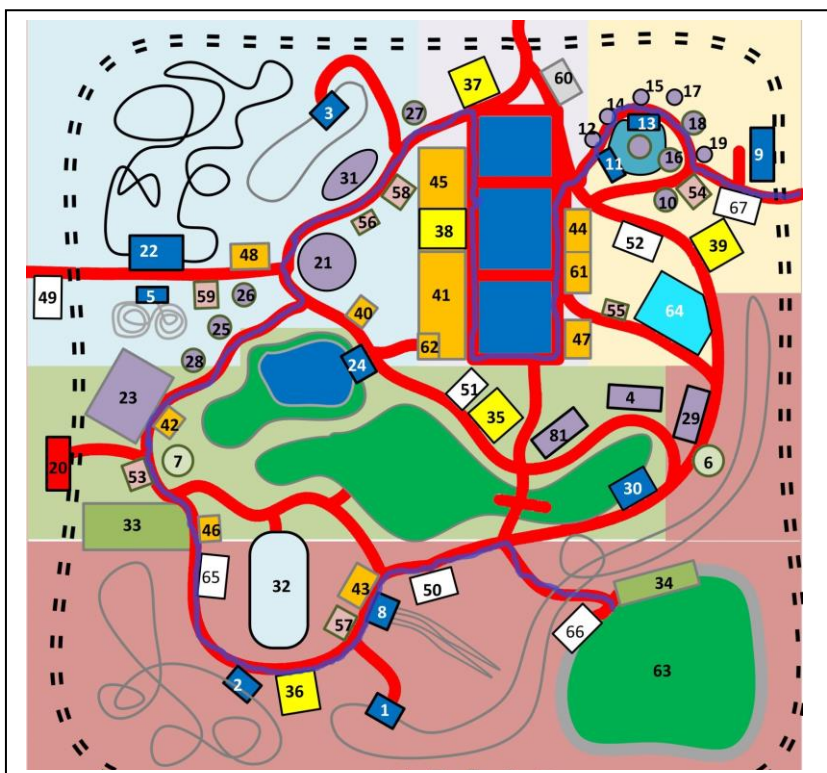
GC 1. Scott is not a paying customer and does not have an ID. Describe Scott Jones' activities in the park during the three-day weekend. Who does he spend most of his time with? When does he arrive? When does he leave? What route does he follow? Limit your response to no more than 10 images and 1000 words.

Scott Jones and his entourage are at the park all three days. Scott himself is not wearing a sensor, but Scott and his entourage are always together. Scott's movements can be tracked by identifying the movements of his entourage, who are tracked. Scott has six appearances scheduled (two each day). He performed in the first five shows, but the sixth was canceled due to concern about Scott's safety.

Scott's entourage consists of the following ID numbers 1080969, 1600469, 1629516, 1781070, 1935406, 521750, and 644885.

For the morning daily show, Scott and his entourage arrive at the park at 8:45 am and leave the park between 11:30 and 12:30. For the afternoon daily show, Scott and his entourage arrive at the park at 13:45 and leave the park at between 16:30 and 17:30.

The route taken by Scott and his entourage is indicated by a blue line on the map below. When the show is finished, Scott and his entourage leave the stage, reversing their path from earlier, departing through the east exit. As he walks through the park, he is stopped by fans for autographs to the extent permitted by his entourage.



GC 2. Identify up to 8 issues with park operations during the three-day weekend. Provide a rationale for your answers. Limit your response to no more than 8 images and 800 words.

The main park operation issues that occur are the closure of the Pavilion (Attraction 32) after the Sunday morning show for Scott Jones and the subsequent cancellation of the afternoon show.

Other park operations issues that may be reported include:

1. Difficulties with the park app. On Saturday, around 15:53, data loss may be reported from some visitors. This happens again on Sunday, around 10:23. There is also a problem with the data reported for visitor 1983765 (this turns out to be the prime suspect). Starting at 20:18 on Saturday, Eddie Smith tampers with his app in a test of disabling the tracking feature. Instead, this creates spurious duplicate entries of movement placing him at a different part of the park. When he enters the park on Sunday employees reinitialize the app since it is “acting funny.”
2. Issues described as part of Mini-Challenge 1.3 may also be reported here.

GC 3. For the crime, describe the following, and provide a rationale for your answer.

- a. **When did the crime occur?**
- b. **Where did the crime take place?**
- c. **Who are the most likely suspects in the crime?**

Limit your response to no more than 5 images and 500 words.

The crime occurred on Sunday, between 9:15 and 11:33 in Location 32, the Creighton Pavilion.

The most likely suspects are ID 1983765 (representing Eddie Smith, the prime suspect) and his accomplices, ID 1089132 and ID 1723967. The prime suspect and the two accomplices performed surveillance of the Creighton Pavilion and the park perimeter and exits on Friday and Saturday.

The Pavilion closed 9:30 – 11:30 and 14:30 – 16:30 every day during the Scott Jones shows. The park was short on security, so they had to close this exhibit in order to provide sufficient security for the Scott Jones show. On Sunday, the prime suspect stayed in the Pavilion while it was closed to the public, while his accomplices remained on watch outside. Once Security came around to re-open the Pavilion, the prime suspect left the Pavilion and immediately left the park.

The vandalism in the Pavilion was discovered by some of first park visitors who went into the Pavilion after its 11:30 re-opening on Sunday. This is indicated by the increase in communications as the park visitors discover the vandalism, report it, and talk about it among their groups and with their friends and family outside the park.