Who disrupted the fun at DinoFun World? VAST Challenge 2015

Anmol Singh Suag, Sanuj Bhatia
Department of Computer Science
University of Massachusetts Amherst

Problem Description

o DinoFun World, an amusement park, is hosting a football star over 3 days from Friday to Sunday

o Scott Jones has 2 shows scheduled everyday, and also a Pavilion to showcase collectibles and archives from his career

o On Sunday, his statue was vandalised and it caused mayhem among the public

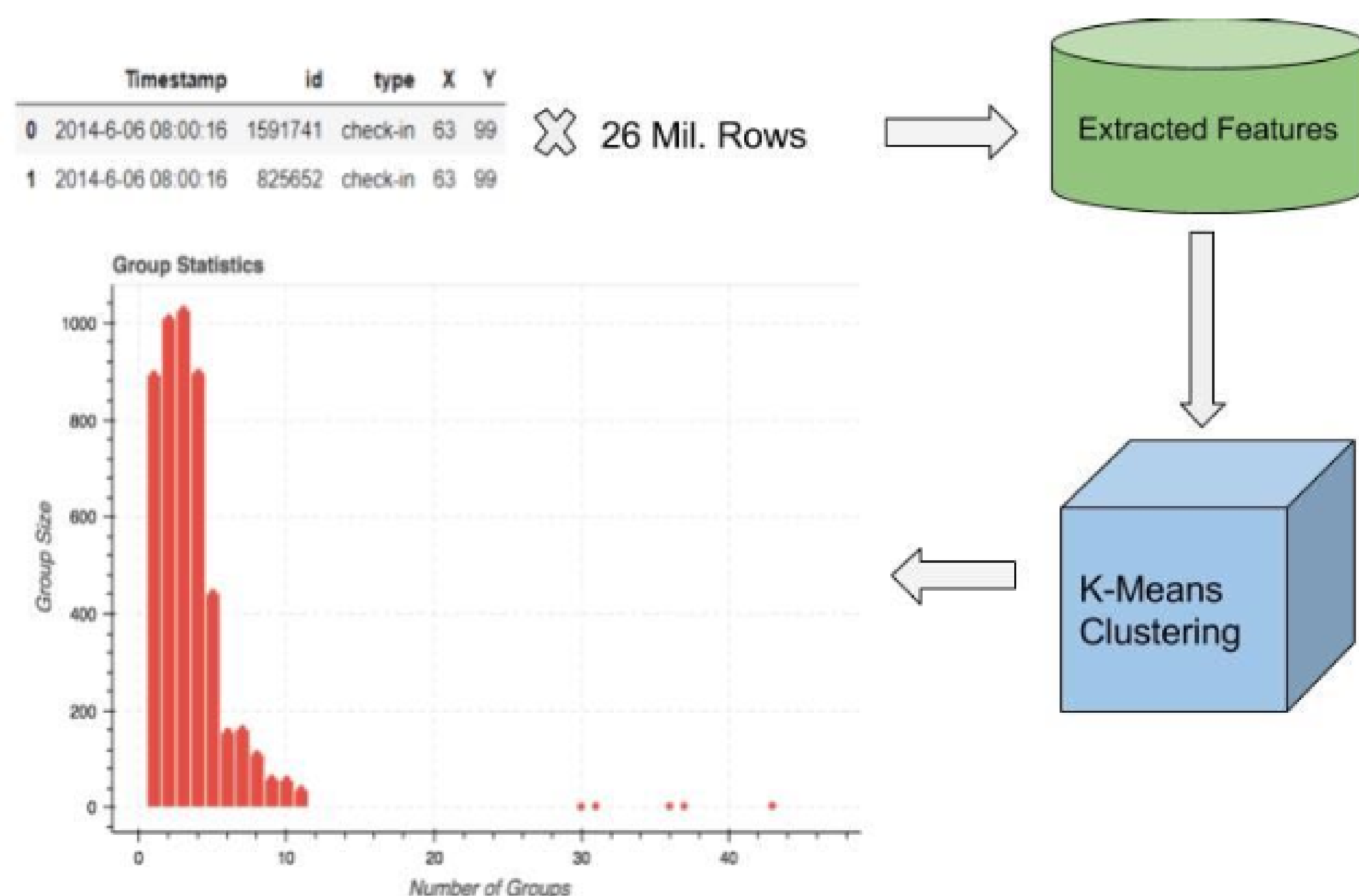o The challenge is to understand what happened and who caused it

This was the VAST Challenge 2015 and consists of 2 Mini Challenges and a Grand Challenge
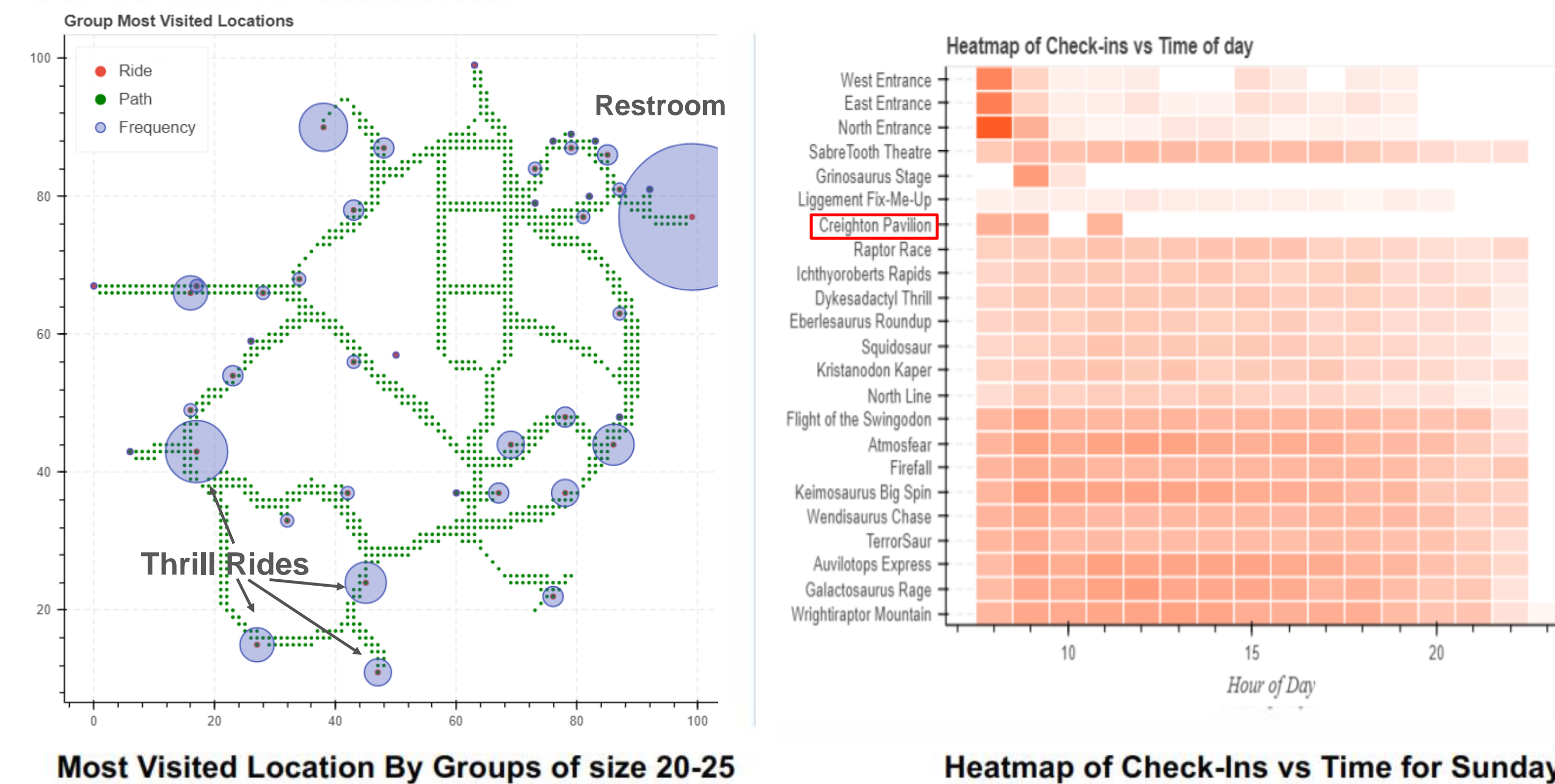
Mini Challenge 1

Dataset

o Visitors' location tracking data is available through the park mobile application.

o The X,Y co-ordinates of all Visitor IDs are available over the 3 days for park movement and attraction check-ins

o Check-in frequencies are used as features to find visitor groups using K-Means. Movement data is used to find paths in the park

Finding Visitor Groups

Observations – MC1

o Groups of size 20 - 25 check-in a lot at thrill rides and also to the rest-rooms

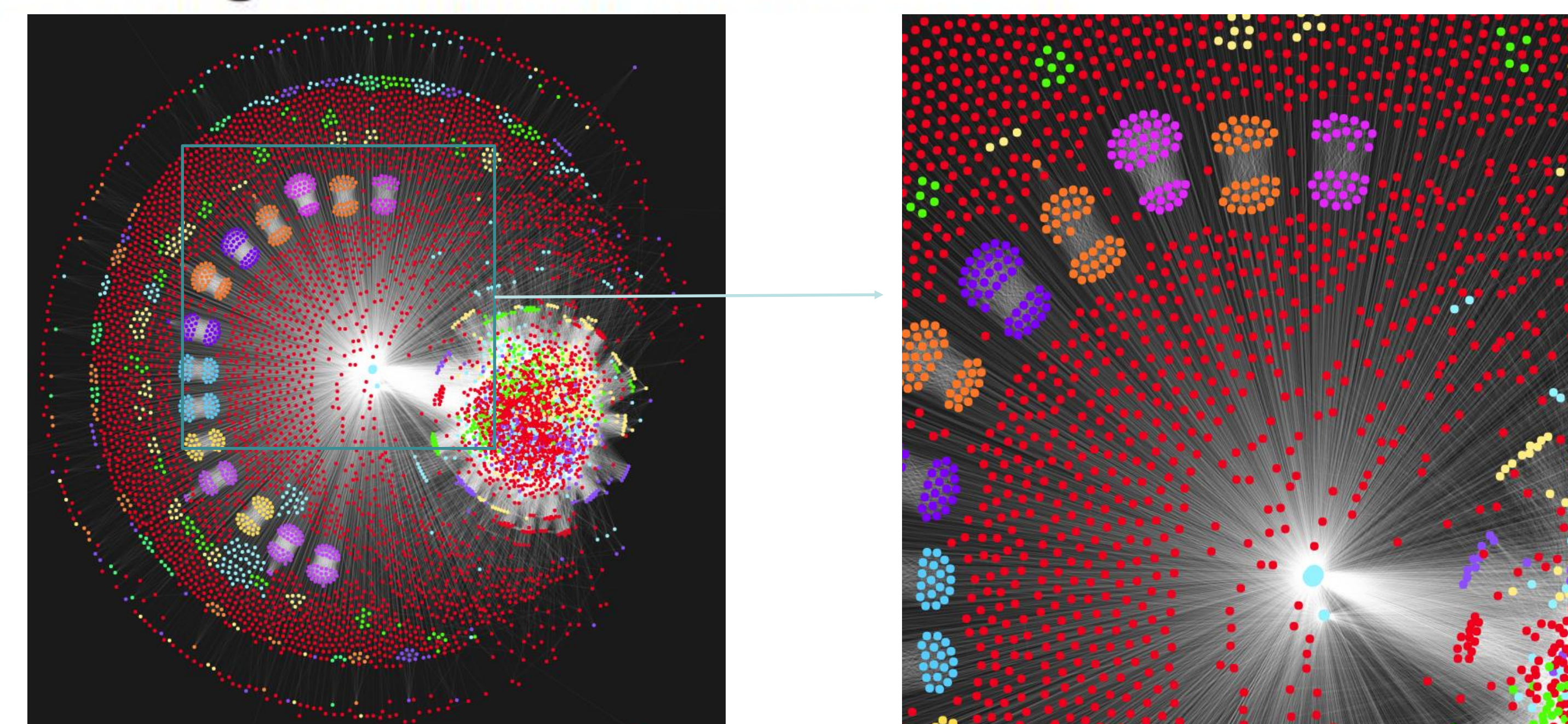o Creighton Pavilion had no check-ins after 12:00 on Sunday - might have been closed due to vandalism

Most Visited Location By Groups of size 20-25

Heatmap of Check-Ins vs Time for Sunday

Mini Challenge 2

Dataset

o Park Application let's visitors send messages to each other and outsiders

o Communication Data for the 3 days of Scott's Weekend are provided

o Data was visualised as a graph to see communication patters and grouped to find IDs that communicate oddly
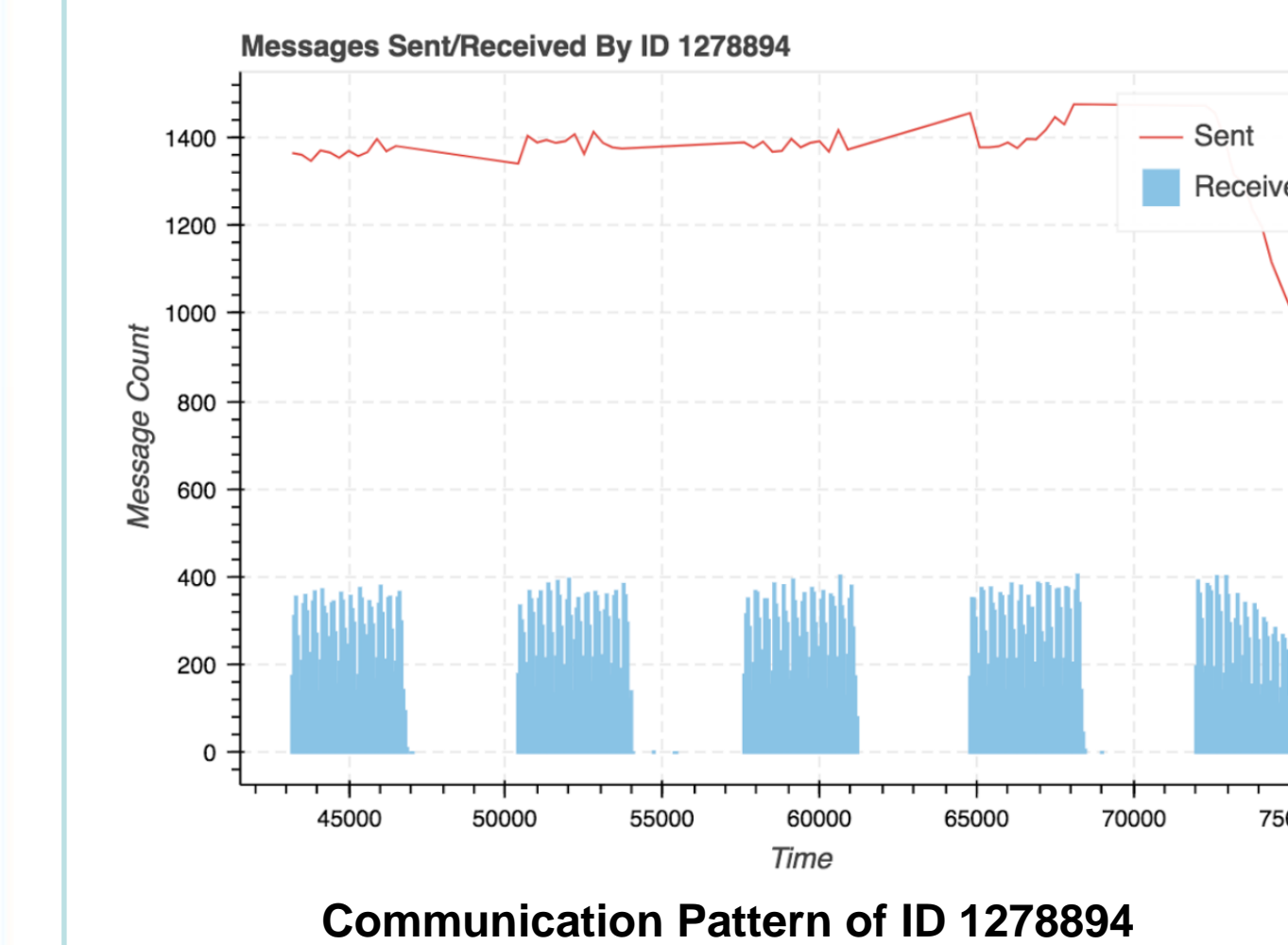
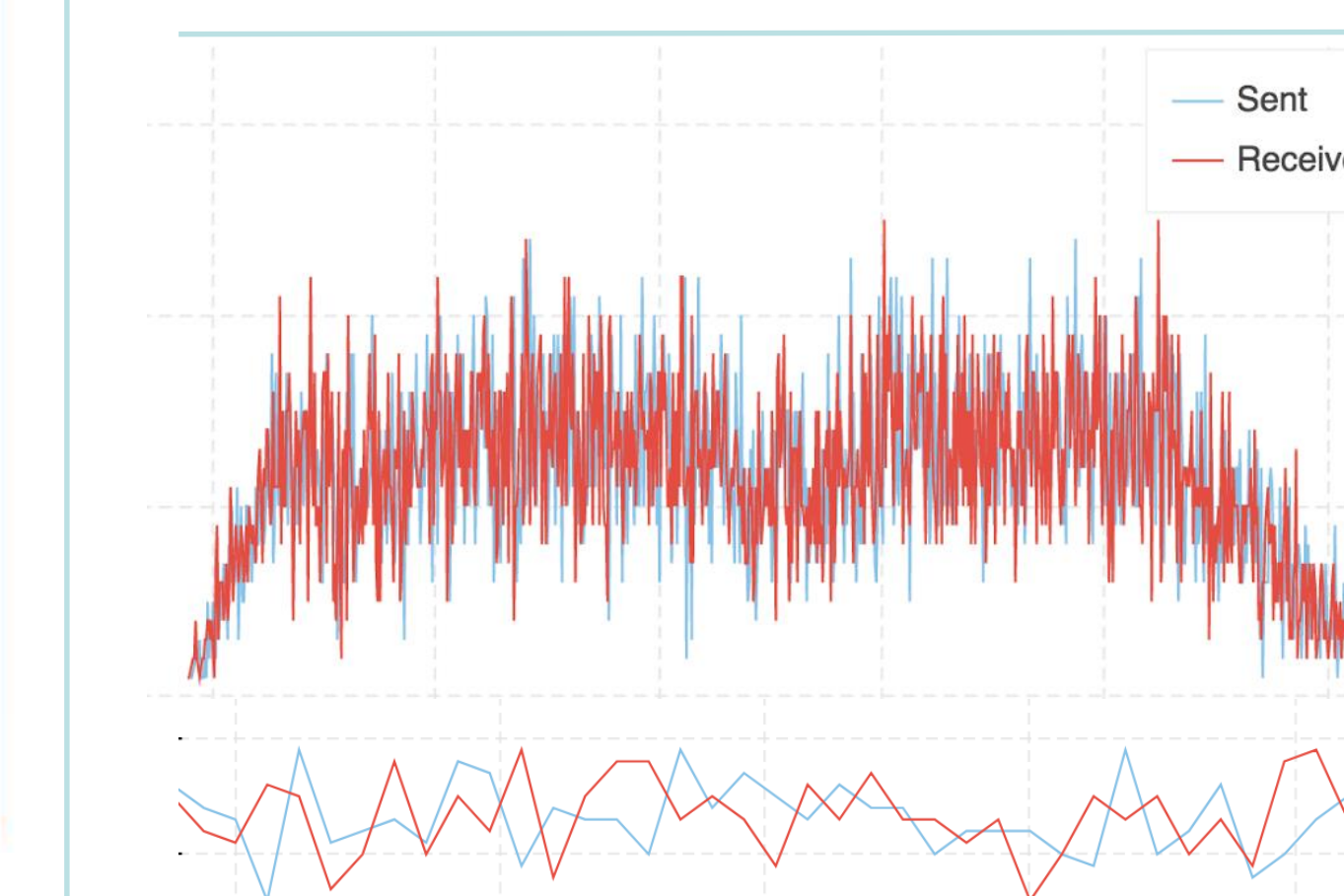### Creating a communication network

With unique IDs as nodes, and text communication between visitors as undirected edges, we imported a **NetworkX graph** representation from bokeh to **Gephi.** The layout is generated using the **Fruchterman Reingold method,** a force-directed algorithm, and then **plotted in Bokeh as an interactive graph.**

• The three central nodes of the graph are **External, ID 839736, and ID 1278894.** These have the most communication to/from them, and the IDs are probably park service IDs.
• The colored components represent groups that have communication edges within themselves only, apart from messages to and from the park service IDs.
• Groups vary in size from 1 to 40. Other interesting patterns also emerge from this graph.
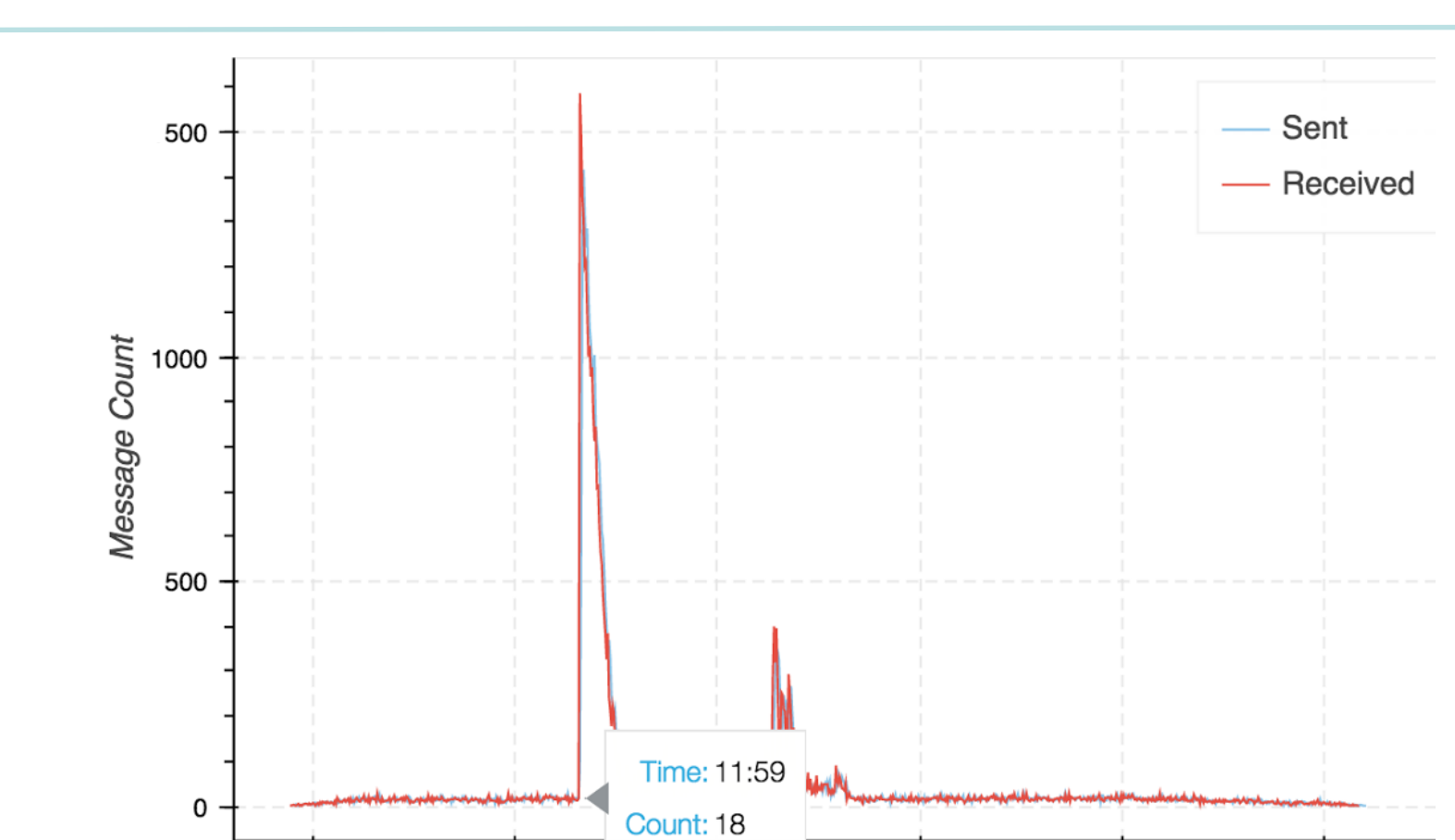
Observations – MC2

Communication Pattern of ID 1278894

• ID 1278894 sends messages to every other ID in the park in the same pattern throughout the 3 days.
• We hypothesise that it is a an automated park advertisement/information system.
• It receives messages **5 times in a day in one hour intervals** which are themselves spaced by an hour. This is possibly a **survey or some game** that opens every other hour and park visitors participate in it.

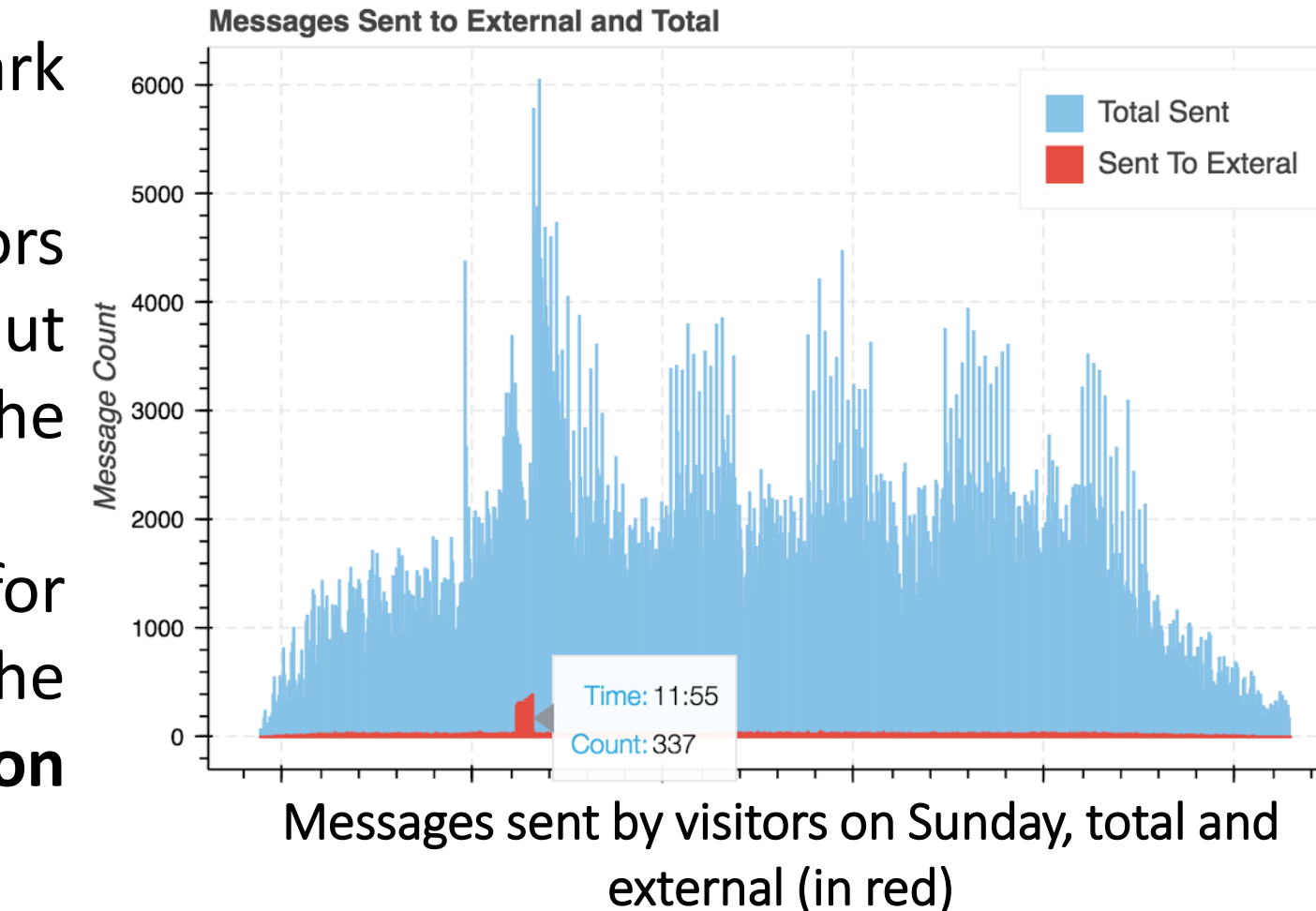**ID 839736 (park helpdesk) sends and receives messages continuously on Friday and Saturday.**

There is a spike in communication to the helpdesk around 12 when visitors discover vandalism

• Messages sent to recipients external to the park spike on **Sunday 11:45 to 12:00.**
• The spike in external messages is because visitors started informing their friends and family about the vandalism as they discovered it in the Creighton Pavilion.
• The **total amount of messages sent has a spike** for that time as well, hence we hypothesize that the vandalism was discovered around **11:45 AM on Sunday, when the Pavilion was reopened.**

Messages sent by visitors on Sunday, total and external (in red)

Grand Challenge

In order to find Scott Jones and his escorts, we filtered the list of **visitors who came to the park everyday** by removing those who **checked in at any rides.** We found **8 IDs** who all have taken the same path around the park and came in everyday before his show starts and left shortly after the show ended.

We think that these 8 IDs are Scott's bodyguards. They have no communication record as they might be communicating via their own walkie-talkies. A map of DinoFun World is show with the movement data of ID 1080969 in blue.

• They enter from the east entrance and walk to location 63, Grinosaurus Stage, where the show is held, twice a day.
• On Sunday, they come in the morning and leave after the show. They do not come again for the afternoon show, as it is now cancelled because of the crime.