Haztrailz: Exploratory Analysis of Trajectory and Sensor Data

Sriram Karthik Badam, Christoph Kinkeldey, Petra Isenberg

Haztrailz: Exploratory Analysis of Trajectory and Sensor Data

VAST 2016 Mini Challenge 2 Award: Honorable Mention for Clear Analysis Strategy

Sriram Karthik Badam*
University of Maryland, College Park

Christoph Kinkeldey†
INRIA Saclay

Petra Isenberg‡
INRIA Saclay

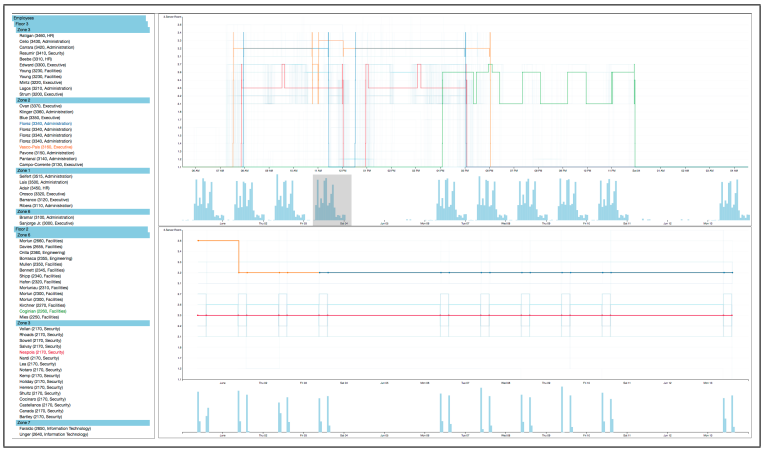

Figure 1: Custom trajectory analysis tool developed for understanding the employee movements within the GASTech building. The tool shows two sets of staff trajectories within the building—one based on their proximity card check-ins (top) and the other based on the mobile robot (bottom). Employee locations (floor, zone) are connected with lines in each trajectory visualization and the employees are organized by department on the left. Bar charts of total check-ins (per hour) over the entire time span are used to select specific time ranges in the visualizations.

Abstract

The mini challenge 2 from VAST 2016 dealt with understanding the operations data from GAStech, a fictional company. We analyzed two weeks of this data including, (1) employee locations collected using proximity cards, and (2) sensor data containing temperatures, heating and cooling status, and chemical concentration levels. Our approach involved data cleaning and consolidation using R, the development of a custom trajectory visualization tool for the analysis of location data, as well as the use of existing analysis tools for the combined analysis of sensor and location data. In this paper, we discuss our analysis process, and report patterns and anomalies, as well as suspicious activities identified within the building.

1 Introduction

The scenario for the 2016 VAST challenge involved GAStech, a growing company in the island of Kronos that transitioned to a new, state-of-the-art building with high energy efficiency standards and security measures. The three-story building is divided into several HVAC (heating, ventilation, and air conditioning) zones with each zone containing many sensors that help track temperature, heating/cooling, and chemical concentrations of CO_2 and Hazium (a fictional gas). The company uses proximity cards to track their employee movements. Furthermore, the company has a smart robotic mail delivery system, *Rosie*, that travels the halls on a specific route and identifies proximity cards using a mobile proximity sensor.

*e-mail: sbadam@umd.edu
†e-mail: christoph.kinkeldey@inria.fr
‡e-mail: petra.isenberg@inria.fr

Given these state-of-art data collection technologies, it is a challenge for GASTech management to make sense of this heterogenous operations data. Our goal is to use visual analytics to identify typical patterns, trends, and anomalies in two weeks of operations data to find any suspicious activities. At the same time, identifying a good strategy to achieve this goal is also of interest. In our analysis, we approached this (mini) challenge by first assembling the necessary infrastructure to understand the different parts of the operations data. Using this infrastructure, we formed three perspectives into the data, (1) understanding typical patterns within the building, (2) interpreting trajectories over time for employees of different departments to observe behavior outside the normal, and (3) exploring sensor data using small multiple charts by highlighting anomalous values. We later connected the generated observations using these perspectives through group discussion to speculate on suspicious behaviors and events. This procedure is popular in collaborative visual analytics where coupled+decoupled sensemaking activities [1, 2] are advantageous for insight generation.

2 Analysis Process

We started the data analysis by first making some grounding assumptions regarding (1) what constitutes a typical pattern and (2) what are the critical levels for certain sensors (e.g., for CO_2 and temperature). At the same time, we processed the data with R, by joining, cleaning, and transforming the data files, and created charts of typical employee behavior. Figure 2 shows an example of the working hours of GASTech employees from different departments. We used QGIS to map the mobile proximity data collected by the robot to a spatial map of the floors in the building. We manually assigned the employee locations in the map with the corresponding proximity zones in each floor (e.g., meeting hall, offices etc.).

We developed a custom trajectory visualization using the D3 toolkit where the locations covered by the employees are connected

First log of the day, Admin, Eng, Exec, HR, Sec

Day of Timestamp

Dep	May 31, 2016	June 1, 2016	June 2, 2016	June 3, 2016	June 6, 2016	June 7, 2016	June 8, 2016	June 9, 2016	June 10, 2016	June 13, 2016
Administration	07:15	07:15	07:15	07:15	07:15	07:15	07:15	07:15	07:15	07:15
Engineering	07:15	07:16	07:15	07:15	07:15	07:15	07:15	07:15	07:15	07:16
Executive	07:30	07:00	07:00	07:14	07:20	07:00	07:14	07:07	07:15	07:15
HR	07:30	07:33	07:45	07:30	07:30	07:29	08:00	07:30	08:00	07:30
Security	07:32	07:30	07:30	07:31	07:31	07:30	07:30	07:31	07:30	07:32

Last log of the day, Admin, Eng, Exec, HR, Sec

Day of Timestamp

Dep	May 31, 2016	June 1, 2016	June 2, 2016	June 3, 2016	June 6, 2016	June 7, 2016	June 8, 2016	June 9, 2016	June 10, 2016	June 13, 2016
Administration	17:48	17:33	17:32	17:32	17:33	17:33	17:32	17:32	17:32	17:32
Engineering	17:21	17:21	17:22	17:22	17:21	17:19	17:21	17:20	17:19	17:21
Executive	17:57	18:01	17:07	18:03	18:02	18:02	17:46	17:45	17:46	17:10
HR	17:32	17:12	17:32	17:32	17:50	17:46	17:37	17:47	17:30	17:51
Security	17:18	17:21	17:20	17:18	17:21	17:21	17:23	17:23	17:19	17:23

Figure 2: Average time of first and last login of the day for sub-groups within the Facilities and Information Technology Departments.

through lines over time (Figure 1) to uncover the trajectories of the employees. For example, the employees enter the building through the main entrance (floor.zone 1.1) between 7:00 and 8:00am and use the elevators or stairs (floor.zone 1.4) to get to their offices. At noon, many have lunch at the Deli on the first floor (floor.zone 1.2) and then return to their offices. In between, it is common that employees move to the zones with meeting rooms such as zone 2.6. Around 17:00 they leave their floor via the elevator / stairs again and the building via the main entrance. Figure 3 shows the typical activity of two employees from engineering and administration respectively. This tool further helped us observe events that deviate from the typical behavior. For example, some employees did not swipe/check their badges when leaving the building.

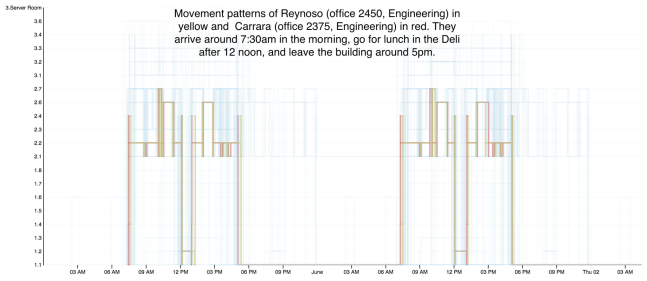

Figure 3: Example movement of Walton Reynoso (yellow) in Engineering (office 2450) and Lise Carrara (red) in Administration (office 3420), against the overall movement patterns of all employees (blue).

To analyze the sensor data, we used Tableau to generate small multiple line charts for each sensor grouped by zone and floor. We started by visualizing the average sensor readings (e.g., heating and cool set points) to see if there are specific days with particularly high or low values. This helped us find that the building seems to have been set to a slightly higher heating (29-30°C) and cooling (26-27°C) temperatures on June 7th and 8th. This also gave an idea of the "normal" value ranges to further search for anomalies in sensor readings. We also found various anomalies in the CO_2 levels and Hazium spikes on certain HVAC zones in the three floors.

As mentioned earlier, we made some grounding assumptions by researching the expected temperature levels in places like Kronos and also the critical limits of CO_2 concentrations. This helped us find identify when the sensor readings are outside the normal and sometimes even affect the staff within the building. For example, CO_2 levels in the dataset were close to health critical limits of 5000ppm on June 7th and 8th. In particular energy zones 14 (meeting room 2700) and 11 (corner offices) were affected. In addition,

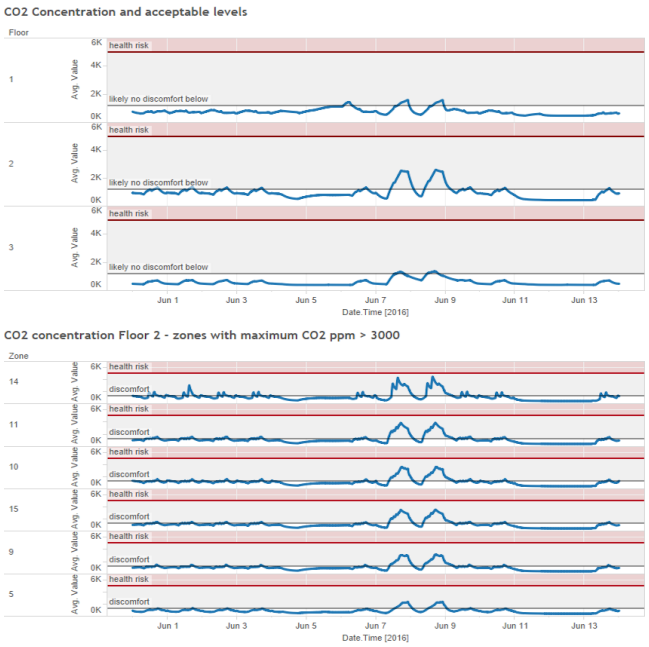

Figure 4: CO_2 concentrations; average for each floor (top image) and the most affected E-zones on Floor 2 (bottom image). The red band indicates ppm rates that pose a health risk while the gray band indicates levels of CO_2 that may be experienced as uncomfortable.

1200ppm can be the limit before people are bothered by body odors in the air. This critical limit was reached on June 7th and 8th on all floors—in particular Floor 2 Zones 14, 11, 10, 15, 9, 5 (Figure 4).

Finally, we combined the observations from the three perspectives to find suspicious activities. A significant finding from this is the likely tampering with the HVAC system by the holder of Patrick Young's badge (pyoung001). The (probably) stolen badge enters the server room on four occasions on June 2, 10:21, June 3, 10:22, June 8, 10:21, and June 10, 10:22. The sensor data for the server room does not show an immediate regular reaction to these entry times, so it is unlikely that pyoung001 just flipped a switch. However, it is possible that he changed future settings (if the server room has an HVAC controller) or tampered with settings elsewhere.

3 Conclusion

We performed an analysis of the operations data including employee proximities and building sensor information using three perspectives. We charted the common trends based on average sensor values and typical movement patterns. We were able to identify unusual movement patterns in the trajectory data using a custom trajectory visualization tool. We also found anomalies in the sensor data through small multiple charts. Finally, we discussed the findings as a group to speculate about suspicious activities in the building. An important outcome of our analysis is the likely tampering of the HVAC system using Patrick Young's badge (pyoung001).

References

[1] P. Isenberg, N. Elmqvist, J. Scholtz, D. Cernea, K.-L. Ma, and H. Hagen. Collaborative visualization: definition, challenges, and research agenda. *Information Visualization*, 10(4):310–326, 2011.

[2] W. McGrath, B. Bowman, D. McCallum, J.-D. Hincapie-Ramos, N. Elmqvist, and P. Irani. Branch-explore-merge: Facilitating real-time revision control in collaborative visual exploration. In *Proceedings of the ACM Conference on Interactive Tabletops and Surfaces*, pages 235–244, 2012.