Visualisation of Building and Human Sensor Data

Wesley Fabrice, Cheryl Lee Mei Hui, Linda Teo Kwee Ang

Abstract - With the rise in security concerns, sensors are increasingly being used in everyday life. This report shows how we can easily and effectively visualize abnormalities present in the building data used in the VAST Challenge 2016 using D3.js and Tableau. Key graphs that were utilized include BiPartite chart and Gantt chart for human movement data and Sparklines for visualization of building sensor data such as temperature, composition of air, or electricity usage.
Index Terms: Visualisation, Sensor, Tableau, D3.js, BiPartite, Gantt

1. INTRODUCTION

Human sensor data and physical sensor data may be fused synergistically for improved coverage and better quality of decisions (Pei-Hsuan Tsai et al, 2014). In Visual Analytics Science and Technology (VAST) Challenge[1] 2016, human sensor and building data of a company, GAStech, is to be analysed. GAStech has been operating a natural gas production site in the island country of Kronos. Its new 3-storey office building is fully instrumented with sensors that identify everything from building temperatures to concentration levels of various chemicals. Due to the security processes of the company, employees are required to wear proximity cards which allowed their movements to be recorded and tracked.

Two weeks' worth of building sensor data and employee proximity data were provided, to be analysed for patterns of concern.

2. MOTIVATION AND OBJECTIVES

The building in this scenario was described as built to the highest energy efficiency standard. This kind of intelligent buildings increases operational and energy efficiency, and at the same time, generates high amount of data which supports real-time activity monitoring. The proximity card records provide another set of data specific to individual employees. Both sets of data may complement each other, and may be used to detect anomalies in each other.

The dataset comprised the following:

- HVAC readings for the entire building
- HVAC readings for each floor
- HVAC readings for specific zones within each floor
- Proximity (movement) records for each proximity card holder.
- Proximity (movement) records captured by the mobile mail delivery robot Rosie who travels the halls periodically, moving between floors in a specially designed chute.

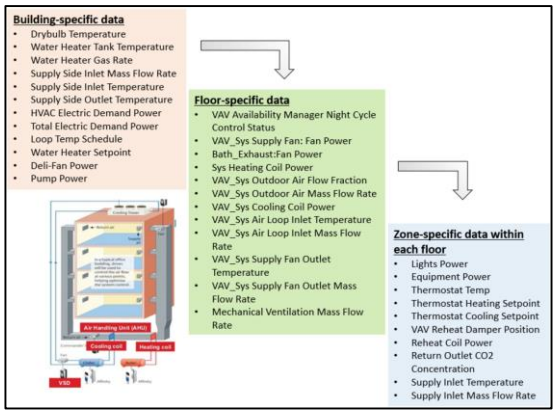

Figure 1: Listing of HVAC variables in the dataset, with an illustration on how the HVAC system works

[1] The Visual Analytics Science and Technology (VAST) Challenge is an annual contest with the goal of advancing the field of visual analytics through competition. 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, as well as an opportunity to advance the field by solving more complex problems.

One situation included in the VAST challenge, is the HVAC readings on concentration levels of various chemicals such as carbon dioxide (CO2) and Hazium. The latter is a recently-discovered and possibly dangerous chemical. Overall, we also aim to design suitable dashboards and visualisations which can aid in the real-time monitoring of both HVAC and proximity records.

3. APPROACH

Our approach comprises of several steps done iteratively.

Analysing the HVAC data

Exploratory analysis was firstly done to comprehend the datasets. The HVAC data is specialised knowledge which a non-technician is unlikely to understand. Certain online research was done to understand what each variable refers to, but it is not adequate to provide an understanding on how each variable interact with one another.

Map shapes were draw using Tableau, based on the office layouts to facilitate **geospatial analysis**. This was necessary because the HVAC readings was categorised in zones. A **temporal analysis** can then be performed based on the readings in each zone.

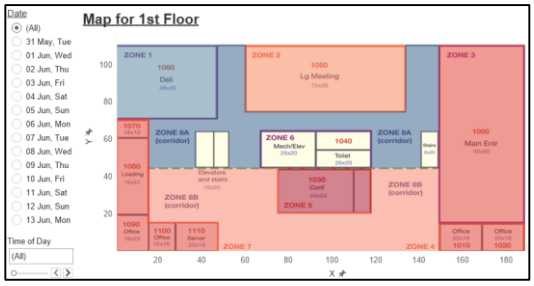

Figure 2: Example of a map visualisation in Tableau.

Analysing the proximity data

The proximity card ID captured in the proximity data were short-forms of the employee names. Hence certain **data cleaning** is required so as to facilitate exploratory analysis on the employees.

There were also inconsistencies in the proximity card ID data, hence they needed to be checked thoroughly. Having matched the proximity card ID to specific employees, the key data on the employee's department and office location could be updated into the dataset.

Similar to HVAC data, the map shapes were drawn using Tableau based on the proximity zones used in the proximity data, to facilitate **geospatial** and **temporal analysis** of the proximity data.

Visualising the data

While the HVAC and proximity data can be explored and analysed via Tableau in mapped layouts and graphs, more appropriate visualisations will be necessary to for more effective analysis.

Different applications and methods were attempted to derive the appropriate visualisations. **Tableau** and **D3.js** were used.

Sparklines

Sparklines give a clear and compact graphical representation of the data. Tableau was used to create the sparklines, using selected HVAC readings like Hazium concentration. This visualisation can be used as a monitor, to see the real-time changes with comparisons of the previous week and average readings.

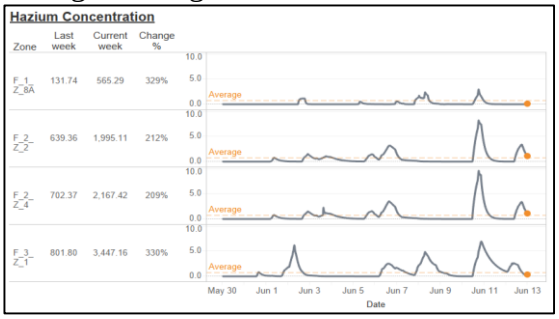

Figure 3: Example of sparklines used to visualise the Hazium concentration

Zoomable area charts

Zoomable area charts were created using D3.js, allowing users to select and zoom in on an area of the chart by clicking and

dragging around the area they want enlarged. This is a good graphical method, given that the HVAC data may need to be examined in granularity, by time of reading by the minutes. For example, by zooming down to specific day frames below, we can see the equivalent expanded version on top with the readings. This zoomable area chart is accompanied with other charts in one webpage so that a few HVAC variables can be visualised at the same time using the same interactive filters.

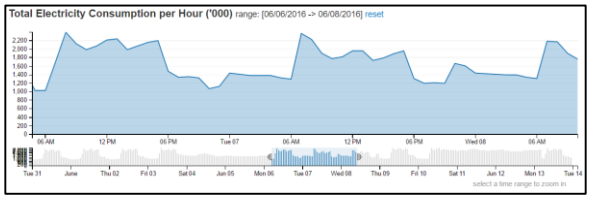

Figure 4: Example of a zoomable chart on the HVAC variable of total electricity consumption

The benefit of using D3,js in this chart, is its light-weight, seamless interactivity with any browser without the need for plug-ins.

Gantt Charts

Gantt charts illustrate timelines and are typically used in project management. It is appropriate as a graphical method in the proximity data, as it is able to reflect the duration of the individual employees at a certain zone at a certain time. For example, the movements of employee Valeria (from Facilities Dept) on 7 Jun 2016 (Mon) can be easily visualised from the following Gantt chart upon applying the filters on day and name.

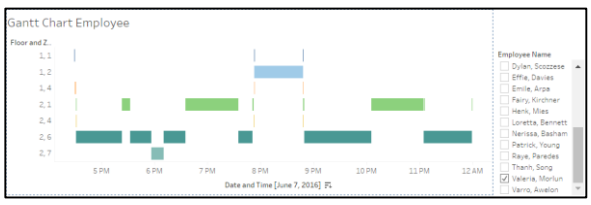

Figure 5: Example of a Gantt chart created using Tableau, on an employee's movements

BiPartite charts

BiPartite charts refer to graphs where the vertex set are divided into two independent sets. D3.js was used to create the BiPartite charts to show the movements of the employees at different zones. Different BiPartite charts were created for the different Depts as it would be too cluttered to stack more than 100 employees on the same chart. The example below is pertaining to the Executives Dept. When mouse-over any of the zones, the vertex could be highlighted to reflect the employees who had movements into that zone as well as their proportions. Like-wise, when mouse-over any of the employees, their corresponding movements into the zones will be highlighted. The chart on the right is pertaining to the employee's working office location, so comparisons can be made on how much movements the employee had in his/her working office vis-à-vis outside working office.

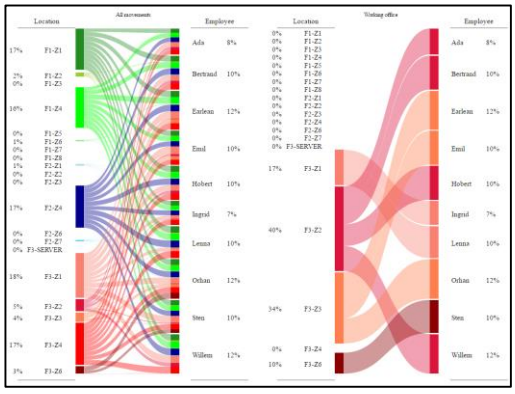

Figure 6: Example of BiPartite chart for Executives Dept

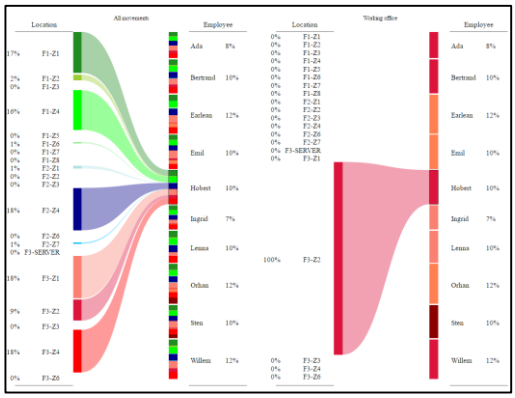

Figure 7: Example of BiPartite chart for employee named Hobert from Executives Dept

4. FINDINGS

Anomalies were identified in the HVAC data by comparing the patterns of the readings with that of their preceding weekday. This is based on the assumption that the first week's operation was under normal conditions. Two key anomalies are as follows:

- **Hazium concentrations** were highest on 11 Jun 2016 (Sat) from 6pm to 8pm, at all 4 zones where Hazium is monitored. The readings were highest at 2nd floor zone 4 which are offices of the Facilities Dept. Fortunately, nobody was present at those locations working during those timings, otherwise there could be health hazards. It was also high on 3 Jun 2016 (Fri) around 7am, at the CEO's office at 3rd floor. It is unclear whether the CEO was present at his office during that time or not, as there were irregularities in his proximity records from 2 Jun 2016 to 3 Jun 2016.
- **CO_2 levels** were higher than 2000 ppm at certain zones between 5 and 8 Jun 2016 (Sun to Wed). Such levels could lead to headaches. This could indicate certain malfunctioning of the HVAC system which is supposed to regulate the level of CO_2 in the building.

The challenge also required patterns to be analysed for the proximity data. Some of the key findings are as follows:

- Three employees, Lise, Mat and Linda from Administration Dept came to office over weekends on 5, 11 and 12 Jun 2016 respectively.
- Cornelia from Administration went to the server room alone on 6 occasions during the 2 weeks. The server room is only accessed by the Facilities and IT Depts, and there are no reasons for Cornelia to go there.
- Movement records were recorded at 12am for CEO Sten on 2 Jun (Thu), and Geneviere (from Administration on 7 Jun 2016 (Tues). Given their usual movement patterns, it is quite unlikely that they would still be in office at this time of the day. There could be irregularities in the proximity sensors/data.
- It is somewhat strange that there are no security employees over the weekends

Other than the above, it has been observed that the HVAC data and the proximity data could be **symmetrical with each other**, if the demarcations for both zones matched. For example, the Deli which is at 1st floor is a zone by itself for both HVAC and proximity data. The lighting and equipment data is reflective of the operations of the Deli, which caters to 3 meals for the employees. For the first week of data (assumed to be under normal conditions), the Supply Inlet Mass Flow Rate also corresponds with the number of employees whose movements showed the same location at the same time frames.

5. LIMITATIONS

There are limitations in the data which hindered fusion of the HVAC and proximity data.

Firstly, the HVAC readings and proximity movement records were tracked by zones in the respective floors. However, the demarcation of the zones in HVAC is different from that of proximity. For example, zone 2 in floor 3 for proximity data is actually a combination of some parts of zone 2 in HVAC data, zones 3, 5, some parts of zone 6, etc. As such, it is very hard to match the HVAC readings directly with the proximity data.

Figure 8: Comparing of zones demarcation for 3rd floor.

Secondly, the security requirements for wearing the proximity cards were not strictly enforced – it was mentioned in the challenge topic that not all staff would diligently wear their cards, and worse, some would misplace their cards, and replacements could be obtained from the security office. This leads to a potential gap in the accuracy and completeness of the proximity data.

Thirdly, the proximity data captured by the mobile mail delivery robot Rosie was not likely to be useful, as she travels in a designed route hence is not able to provide comprehensive data.

6. RECOMMENDATIONS

The value of the HVAC and proximity data can be maximised through real-time monitoring, so that anomalies can be quickly detected and addressed. The following are recommended for GAStech's considerations:

- Dashboards can be used to track the HVAC and proximity data
- Zones for HVAC and proximity sensors should be synchronised for data fusion
- Shifts to be added for security employees over weekends
- Investigations to be done on the anomalies highlighted above

7. CONCLUSION

With an increasing number of buildings over the world being equipped with smart sensors capable of tracking movement and environmental changes, visualization techniques and methods suggested in this report provides an example of how such data can be monitored and utilized effectively without much effort.

REFERENCES

A Framework for Fusion of Human Sensor and Physical Sensor Data
Pei-Hsuan Tsai, Edward T-H ; Ying-Jun Lin, Jane W. S. ; Yi-Zong Ou, Jane W. S. ; Chu, Jane W. S. ; Liu, Jane W. S.
IEEE Transactions on Systems, Man, and Cybernetics: Systems, Sept. 2014, Vol.44(9), pp.1248-1261 [Peer Reviewed Journal]

Fault Detection and Diagnosis in Building HVAC Systems
Najafi, Massieh 2010

An ontology framework for intelligent sensor-based building monitoring
Dibley, Michael ; Li, Haijiang ; Rezgui, Yacine ; Miles, John
Automation in Construction, December 2012, Vol.28, pp.1-14 [Peer Reviewed Journal]