THREE-DIMENSIONAL ENABLEMENT OF PLACE-BASED, PANDEMIC BEHAVIORS

S. Bagul[1], D. Laefer[1,2*]
[1]Center for Urban Science + Progress, New York University, USA - ss15624@nyu.edu
[2]Dept. Civil and Urban Engineering, Tandon School of Engineering, New York University, USA - debra.laefer@nyu.edu

Commission IV, WG IV/9

KEY WORDS: COVID-19, three-dimensional epidemiology, Potree, Geospatial, LiDAR, Egress Behavior, Citizen Science

ABSTRACT:
Harvesting usable and meaningful disaster-related, spatio-temporal data at a highly granular level poses major challenges in its cleaning and aggregation. This paper presents a strategy related to those challenges with respect to individual behavior near COVID-19 laden healthcare facilities. This is done to enable the visualizing of egress behavior data as interactive, three-dimensional (3D) scenes to investigate human behavior patterns regarding touch-based, disease transmission. Therefore, the aim is to demonstrate how this concept of 3D epidemiology may provide new mechanisms to understand the relative risk and exposure prevalence for data analysis. This paper demonstrates 3D enablement of disaster-related field data through use of first-hand observations of 1,936 individuals egressing New York City healthcare facilities during the onset of COVID-19 in the Spring of 2020. The observations capture egress behavior in terms of where people go (e.g. coffee shop, Subway) and how they physically interact with the surroundings (i.e. what they touch and how long they remain). This paper introduces a mechanism for automated extraction and 3D visualization of such data in Potree, an open-source Web Graphics Library (WebGL) point cloud viewer. Distinctive vertex shaders are used to distinguish specific destination selection and behavioral patterns (e.g. personal protective equipment usage). Two-dimensional heatmaps are paired with 3D scenes to demonstrate the potential of using 3D visualization of spatio-temporal patterns for visualizing disease transmission potential.

1. INTRODUCTION

For more than a decade, citizen science and crowd sourcing have selectively contributed to understanding major disaster-related events. These ranges from extremely centralized, national undertakings such as the CoCORAHS rain gage project that has more than 20,000 participants in North America (Reges et al. 2016) to the third-party scraping of social media posts (Palen and Hughes, 2018). Furthermore, researchers such as Buytaert et al. (2016*)* have long-argued for the role that citizen science can play in understanding long-term, disaster-related events. While great strides have been made in identifying and scraping (or otherwise collecting) such data, usability remains challenging due to the ad hoc and highly variable nature of the reporting methods and their frequency and continuity, as well as the tremendous amount of noise in the data. These common characteristics of both Citizen Science data and, to an even greater extent, crowd-sourced data pose fundamental challenges in deriving meaningful and actionable information from such data in a timely manner. This paper demonstrates a workflow that was devised for overcoming some of these via the three-dimensional enablement of such a data set.

2. BACKGROUND

The majority of the world's population live and work in urban areas in densities of approximately 4,600 people per square kilometer (OECD 2020). Such high concentrations of people and activities make understanding their mobility and behavioral patterns important for the mitigation of natural and human disasters such COVID-19. Despite the availability of unprecedented levels of data in terms of footfall signals from cell phones and closed caption television footage, much remains unknown as to the trends in mobility and behavioral data that could inform public health and disaster response policy making. Furthermore, the vast majority of work in this area has been confined to two-dimensional (2D) representations, despite the inherently three-dimensional (3D) nature of the activities. To help bridge this dimensionality gap, this paper introduces a mechanism for more nuanced understanding of individual and group behaviors by contextualizing those actions within real 3D spaces. Specifically, this paper enables the distinguishing of specific egress activities through in-situ, 3D scene visualization. This paper demonstrates means to visualize individual movement-based behaviors as a tool to improve understanding of human behavior in future disasters.

The simulation of events has long assisted experts to prepare for events that are not fully knowable. For example, architects use the earthquake simulators to test the building design before starting the actual site work (Rodolfo et al., 2021) and space agencies use the simulators to adapt to the environment on planetary bodies (Mohanty et al.,2020). Similarly, many government agencies and businesses have started using predictions and simulations using historical data. Geographic information systems (GISs) have played an important role as a tool in the management of various disasters such as the 2013 Ebola Outbreak (Selen et al., 2013), as well as Hurricanes Katrina and Rita in 2005 (Li Wei et al., 2008). Over the past decades, leading health organizations such as World Health Organization (WHO), Center for Disease Control and Prevention (CDC) have consistently relied on spatial analysis to manage disease outbreaks and to mitigate their after effects. Mapping events in real-time space enables unprecedented mean to analyze data. To date, this work has been predominantly 2D, despite the ready availability of 3D scene visualization software. This paper pushes the boundary of 3D epidemiology by demonstrating the emerging capabilities of simulating COVID-19 related behaviors at a hyper-local scale directly within the built environment.

3. METHODOLOGY

Scope: This paper demonstrates the potential usefulness of a method to visualize distinctive spatial data of individuals and groups in a fully 3D context. Two-dimensional views hinder an individual's perception and do not support the experiential perception of walking through the city streets. In order to combat these limitations, recent advances in visualization techniques enable more interactivity using 3D environments. Three-dimensional scenes help to blur the lines between the virtual and physical world, arguably increasing an individual's understanding of data. This is demonstrated using data affiliated with a group of five highly divergent New York City health care facilities. Specifically, this paper employs 1,936 firsthand

* Corresponding author

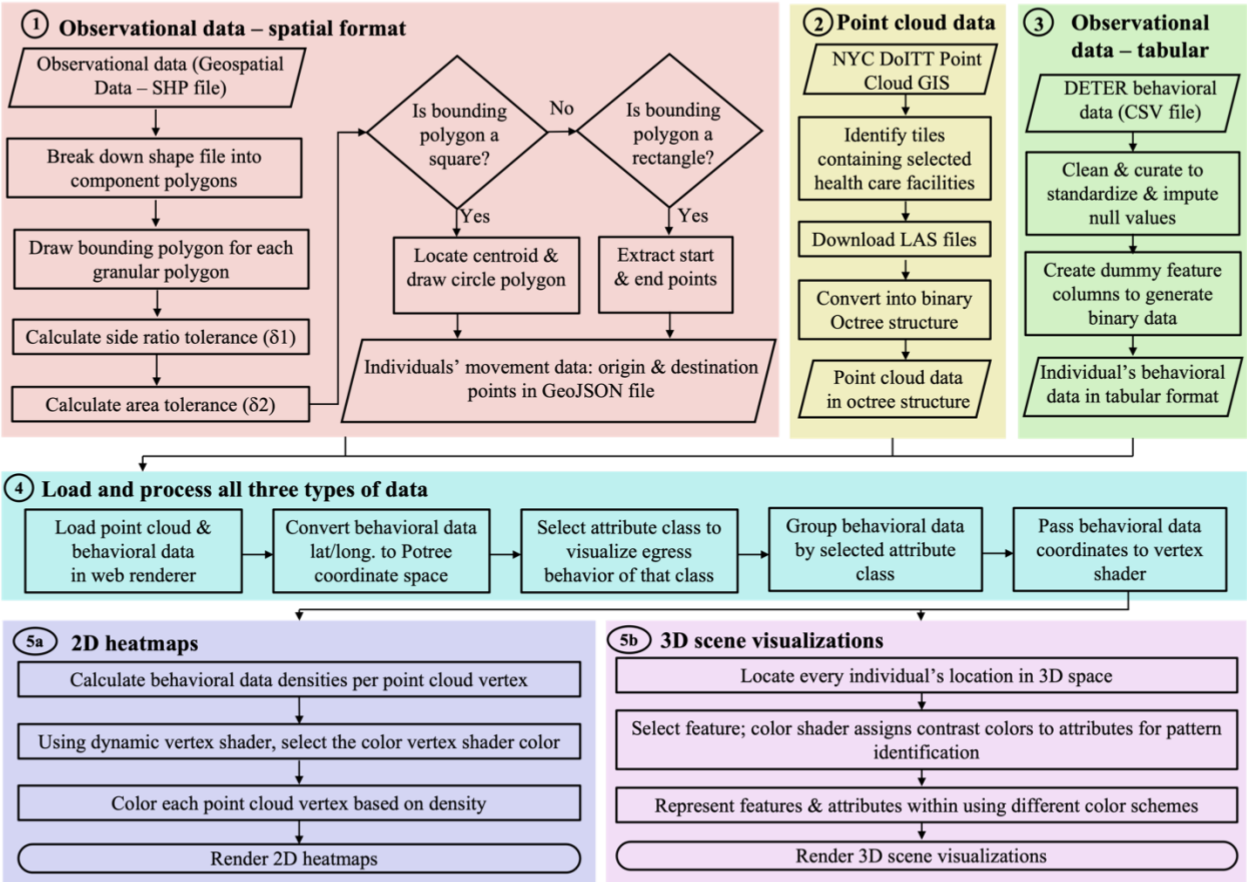


Figure 1. Flow chart of proposed methodology (Sections 2 and 4 adapted from Carey et al., 2021)

observations recorded in the early weeks of COVID-19 in the Spring of 2020 and is a subset of a larger data set of 19 facilities (Laefer et al. 2020), as described in Section 3.1.

The proposed methodology has five major steps (Fig. 1). Step 1 focuses on extracting and standardizing the spatial data from the observational records (stored in a GIS format) [Fig. 1, Step 1]. Step 2 relates to the conversion of medium-density (~15 points/m^2) aerial laser scan tiles obtained from a publicly accessible source (NYC DoITT GIS Unit, 2017) into an octree bin structure using a Potree converter for ultimate use in a Potree-based dashboard [Fig. 1, Step 2]. Step 3 involves the data cleaning of tabular meta-data [Fig. 1, Step 3]. Step 4 describes data loading, data conversion, and attribute selection [Fig. 1, Step 4]. Step 5 presents a pair of visualization options: 2D heatmaps to visualize the high-risk zones and 3D scene visualizations to demonstrate behavioral patterns of individuals [Fig. 1, Steps 5a and 5b].

3.1 Data

3.1.1 Data inputs: The proposed methods employ 3 types of data sets: (1) spatial pathway data of 1,936 individuals egressing select New York City healthcare facilities obtained from firsthand observations and referred to herein as the National Science Foundation (NSF) DETER data (Laefer et al. 2020), (2) publicly available aerial point cloud data and (3) the affiliated anonymized details about individuals such as touch objects, origin-destination location names, and presenting gender in 2D tabular format.

3.1.2 Details about data: The COVID-19 related egress data was collected from March 30, 2020 to May 7, 2020 at 19 locations in 4 of New York City's 5 boroughs with the intent of understanding how people behaved and where they went after leaving COVID-19 laden healthcare facilities (as described in Laefer et al., 2021). The goal was to collect perishable data that could provide insights into pandemic-period actions across dis-disparate communities and distinctive built environments. The resulting 5,065 complete records noted time, date, and location of the start of every record, presenting gender, touch behavior (including cell phone usage), transportation selection, and destination choices (interim and final). These records were collected by 16 unique observers. Out of these data, a subset of 1,936 records were used in this research. Those records were collected at 3 locations in the Bronx (Montefiore Hospital, CityMD Pelham Parkway Urgent Care, Parkchester CityMD Urgent Care) and 2 locations in Brooklyn (Flatbush - CityMD Urgent Care, NYU Langone Hospital Brooklyn) [Table 1]. Additionally, aerial point cloud data of ~15 points/ m^2 from a 2017 flyover of NYC were downloaded from the NYC Open Data portal (NYC DoITT GIS Unit, 2017).

Data Type	Source	Original format	Format used for research
Behavioral data	Laefer et al. 2020	CSV file	CSV file after data curation
Spatial pathway data	Laefer et al. 2020	KML file	GeoJSON
Ariel point cloud data	NYC DoITT GIS Unit, 2017	LAS file	Octree Bin Structure

Table 1. Data types and their sources

All of the original COVID-related data were collected via one of two smartphone apps (either DrawMaps or MyMaps). The observers traced the detailed paths taken by the subjects and marked stopping locations, as well as point of interest (POI) destinations and touch behaviors. The notes and timestamps of the records were manually scraped into a CSV file. The pathway records were then exported into KML or KMZ files and standardized (see Step 3). These were then reunited with the CSV files to generate a publicly accessible data set in a unified GIS format (Laefer et al. 2020).

To depict the behavioral pattern around facilities, origin and destination points needed to be considered. However due to variations in recording styles and formats, of the 1,936 records considered, only 315 destination points were marked explicitly (using circles, ovals or stars). In all other cases, the observers implicitly noted these through the start and end of line segments. By using the extraction algorithm (Algorithm 1), the origin and destination points were extracted explicitly from pathways (lines). The newly formed spatial data resulted in a total of 3,385 start and end location points (315 existing + 3,070 created). The newly identified start and end points were denoted in the metadata by adding the flag value (is_created = 1).

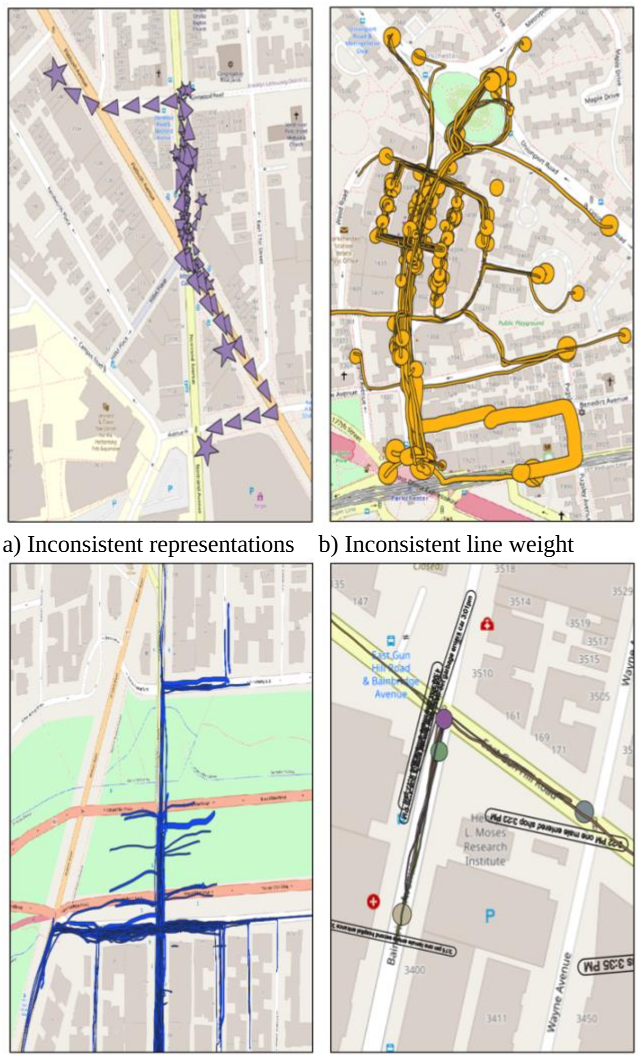
Algorithm 1: Spatial Data Extraction

```
Data:
D (Egress behavior events)

Result:
H (list of potential origin destination points)

function EXTRACT_ORIGIN_DESTINATION_POINTS(D)
    D1 <- POLYGONSPLIT(D)
    δ1 <- SIDE RATIO TOLERANCE
    δ2 <- AREA TOLERANCE
    for each d ∈ D1 do
            bp <- DRAWBOUNDINGPOLYGON(d)
            length <- bp.LENGTH
            width <- bp.WIDTH
            Area <- bp.BOUNDEDAREA
            if length ⊆ [width - δ1, width + δ1] then
                    if width * length ⊆ [Area - δ2, Area + δ2] then
                            H <- H + d
                    end if
            else
                    H <- H + ENDPOINTS(d)
            end if
    end for
    H <- REMOVE_DUPLICATES(H)
    return H
end function
```

The goal was to pair this with aerial Light Detection and Ranging (LiDAR) data. LiDAR data are collected using pulsed lasers to map a 3D environment. LiDAR's use of light allows it to map the environment accurately and more quickly than other approaches. Game development platforms like Unity mimic the real environment by creating virtual objects in games, but LiDAR data are captured from real neighborhoods and city streets. For viewing the LiDAR data, this study starts with the Potree viewer as it is a free, open-source, WebGL-based point cloud renderer. Potree viewer can be used to view billion-point data sets at a high-resolution and with high positional accuracy. The behavioral data (e.g. touch objects, origin-destination location names) and spatial pathway data were collected using mobile apps by drawing the lines of individual's movement (Laefer et al., 2020). This study's aimed to create mechanism in which to better investigate hotspots of COVID 19 spread, with a focus on communal touch locations (e.g. touching the same door handle, hand rail, light pole etc.). From the spatial pathway data, only origin and destination points were extracted and used as hotspot areas for this current stage of the research, as described in the following subsections.

3.1.3 Spatial data preparation: Because of the rapid nature of the COVID research deployment, some of the data collection details were decided on a per observer basis, which was further complicated by the fact that the preferred app was not available for iPhones. Consequently, the spatial data contained various markings with significantly different labels to mark pathways, as well as origin and destination points. To visualize the patterns and to compare the data across healthcare facilities, data was standardized. To standardize the spatial data, all the pathways (lines) and destination points were extracted and separate them from text and the various markers, as described in the data issues section.

3.1.4. Spatial Data Issues: Each marked point in the shape file is a polygon. The original shape file contained polygons corresponding to various shapes - stars, triangles, only lines without origin or destination points, circles of different radius, and text in the form of alphanumeric characters along with spatial data (Fig. 2). Additionally, the majority of records lacked explicit route starting and ending points.



a) Inconsistent representations b) Inconsistent line weight

c) Missing hotspots d) Text overlapping pathways

Figure 2. Record inconsistencies

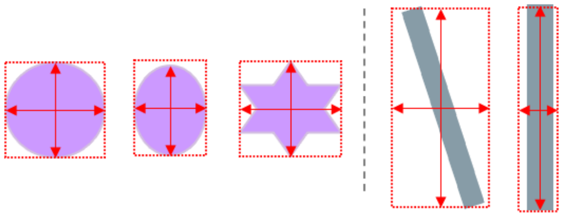
This contribution has been peer-reviewed. The double-blind peer-review was conducted on the basis of the full paper.
https://doi.org/10.5194/isprs-annals-X-4-W2-2022-21-2022 | © Author(s) 2022. CC BY 4.0 License.

If only the original, cleanly marked points were used, there would have been insufficient data points for route mapping, so additional origin and destination points were created, as will be described below.

The records had four major problems. The first was inconsistent origin and destination point marking, as described and clearly shown in Fig. 2a. The second problem was inconsistent thickness of connecting paths in the data, as a direct result of different apps, disparate app settings, and distinctive screen pressure (even within a single record, as shown in Fig. 2b). Next was the problem of missing hotspots, for 1,936 individuals, only 315 points were marked by observers, the rest of the data were not marked properly and only the individual's movement was registered using line shaped polygons (Fig. 2c). The final problem in data extraction was the presence of text in the shape file. Instead of using the separate text editor, some observers wrote the text directly in the shape file. In this case, each letter is considered as a polygon and interferes with useful spatial data. For example, during extraction, the letter 'O' was often confused with an origin or destination point (circular polygon) (see Fig. 2d).

3.1.5 Spatial data extraction algorithm: An extraction algorithm was required for making the data usable and consistent, as the size of the data set and its geospatial nature made manual extraction unfeasible. Specifically, there were nearly 2,000 records, and the representation in the app was image-based, thus not inherently geospatial. So while the shapes indicating paths and destinations had implicit geospatial locations, such information was not available in an explicit way that could be directly derivable. The proposed algorithm followed the steps outlined below.

First, each polygon in the shape file was decomposed into individual components that represent the movement path, origin-destination point, and individual letters of text data marked in the shape file. Next, the shapes were classified. For this, each and every polygon (e.g each star, oval, line) was wrapped in a bounding polygon (Fig. 3). Bounding polygons were defined by using the lower and upper bound coordinates. This enabled their transformation into quadrilaterals (either square-like or elongated rectangles). The square-like shapes typically came from stars, circles, and ovals (Fig. 3a) and represented POIs, while the elongated rectangles were derived from lines of various thicknesses (Fig. 3b) and represented movement pathways. In Algorithm 1, there are two variances used to decide whether or not a shape is squarish more rectangular. Since perfect squares would obviously have a length to width ratio of 1, anything that outside of this by +/- 20% is considered to be more of a rectangle. As a secondary check, there is an area tolerance of +/- 36% of the anticipated area of just the length times the width. This secondary check informs as to the location assignment of start and end points.



a) Rounded shapes   b) Elongated shapes

Figure 3. Rounded shapes (length ~ width) for hotspots and elongated shapes for (length >> width) for movement trajectories

These proposed tolerance thresholds were derived through experimentation on spatial data across the five, selected healthcare facilities. For irregular sized markings, the centroid of the square shaped bounding polygon was used to redraw a uniform shaped circle, which imposed consistency on all extracted hotspots. The center of the circle was used to demarcate the final POI location. In contrast, each elongated rectangle was marked as a line. Each line was the pathway followed by an individual to reach an interim or final destination. To mark the start and end points (origin-destination points), the shorter side of the bounding box was selected. The center points of both the shorter sides (width of bounding box) were marked with points which were then recognized as origin and destination points.

After data cleaning and the imputation of origin-destination points, the usable data were extracted. First the polygon geometry was converted into multi-polygon geometry to visualize in Potree (as per Carey et al., 2021). In that approach, a polygon is a planar surface defined by 1 exterior boundary and 0 or more interior boundaries. Each interior boundary defines a hole in the polygon. A multi-polygon is a multi-surface whose elements are polygons. The Potree viewer works only on multi-polygon geometry (each origin-destination point is marked as a group of points), because Potree in itself is a point cloud format visualization tool.

Potree is a WebGL based renderer, which is used to view point cloud data. Thus, in Potree a line consists of only a set of points, and each point in spatial data represented in form of polygon. For this study, origin- and destination- based data were used to view the hotspot areas rather than the movement paths of individuals.

3.1.6. Point cloud data preparation and extraction: The data POIs extracted from the algorithm were saved to a GeoJSON file. The Potree dashboard requires GeoJSON POI data and the 3D point cloud data to render the visualization. Aerial point cloud data of ~15 points/ m^2 from a 2017 flyover of NYC were downloaded from the NYC Open Data portal (NYC DoITT GIS Unit, 2017) [Table 2]. These were used to provide a 3D contextualization around the five, selected healthcare facilities.

Health Care Facility Name	Latitude, Longitude	LiDAR Points (#)
Montefiore Hospital	40.880, -73.879	10,086,975
CityMD Pelham Pkwy UC	40.855, -73.867	9,989,032
Parkchester CityMD UC	40.837, -73.860	9,634,714
Flatbush - CityMD UC	40.636, -73.946	8,837,017
NYU Langone Hosp. Bkyn	40.646, -74.020	17,067,742

Table 2. Selected healthcare facilities and corresponding LiDAR (UC=Urgent Care)

The tile encompassing each site's location was selected using the provided web interface and downloaded in a LAS format. Each tile was 1000m x 1000m and ranged from 8 million to 15 million points. For the selected facilities, a maximum of 4 tiles were needed depending upon the facility's location with respect to the position in the tile(s). The LAS file was then converted into a binary octree structure using the Potree-Converter tool for easy rendering in Potree Renderer (Potree Development Team, 2021). Based on the findings of Carey et al. (2021) a PC with an Intel i7-7700K 4.2 GHz CPU and Nvidia GeForce GTX 1080 Ti GPU was used for decreasing the input lag encountered during interactive navigation performance; improving algorithm performance for less performant machines is outside the scope of this research. To make the visualization sensible, the elevation of GeoJSON points were set according to the facility location elevation.

This study used a clustering technique to reveal the differentiation patterns in the large dataset like male/female or medical personnel/patients. The data are sliced and used corresponding to each facility to achieve the detailed view. As the focus of this paper is the enablement, limited illustrations and no formal analysis of the data provided.

3.1.8 Loading the data and visualizations: After standardization and preparation of all three types of data (previously summarized in Table 1), all of the relevant data were loaded simultaneously into the Potree dashboard on a site-by-site basis. The Potree dashboard, which is used for viewing the GeoJSON file data and octree structure file data, is interactive and provides the flexibility to visualize the data in two formats, 2D heat maps and 3D scene visualization.

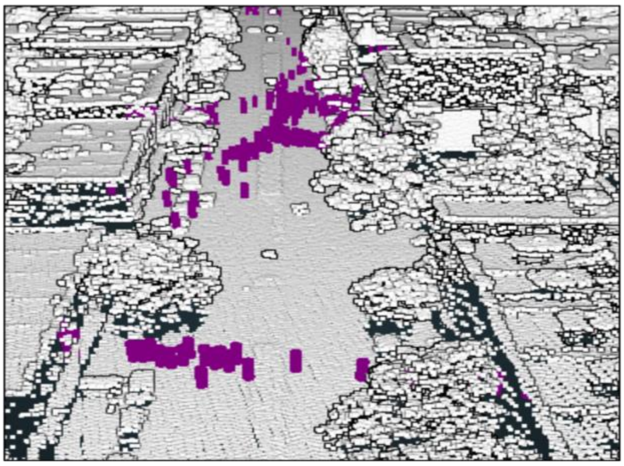


Figure 4. Univariate 3D scene visualization depicting individual egress behavior at a single healthcare facility

In both types of depictions univariate representation is used for visualizing the variations of a single feature (Fig. 4), whereas the bivariate is used to compare the two attribute values across a feature [e.g. usage of personal protective equipment (PPE)]. In bivariate heatmaps (Fig. 5), two values across each feature are visualized by assigning a different color for each of the two values in the vertex shader. In order to make the plots more readable and for distinguishing between different entities, dynamic vertex color shading was implemented.

The method used to visualize the 3D scenes was the cylindrical polygon plots which were created using three.js JavaScript library (*Danchilla et al. 2012*). Geometry used for defining the individuals was "Cylinder Geometry" in three.js library. The 3D scenes were visualized (Figs. 6-7) to know the actual locations of an individual's interactions with the outside environment after egressing a healthcare facility. In the 3D scene visualizations, the individuals were represented using vertical cylinders and were visualized either in univariate or bivariate formats. The univariate plot visualizes the overall view of individuals outside the healthcare facility in the 3D environment. In the bi-variate plots individuals of specific categories (e.g. men versus women; health care workers versus non-health care workers) are distinguished by color (e.g. in Fig. 6, the red cylinders represent the women and blue cylinders represent men).

In each of the illustrations within Fig. 5, distinctive destination choices can be seen. To understand these fully, they need to be examined in detail and compared across all of the facilities within the study set.

Figure 5. Hotspot heatmap for feature (a) "Gender" Male in blue and female in red (b) "PPE Usage" Individuals wearing PPE in green and not wearing PPE in orange (c) "Time Type" Working hour visitors in magenta and off-work hour visitors in cyan (d) "Medical Personnel" Medical personnel in green – others in red

4. RESULTS

This research paper proposed the method to create points representing the origin and destination of individuals which increased the scope of study to analyze the high-risk zones at five health care facilities in New York city (Table 1). The inclusion of multiple locations and 3,000+ data points led to high intensity heat maps across locations which made the comparative analysis possible. Arguably, the comparative analysis of egress behaviors of individuals at various health care facilities in New York City can lead to a better understanding of the spread of epidemic in urban areas.

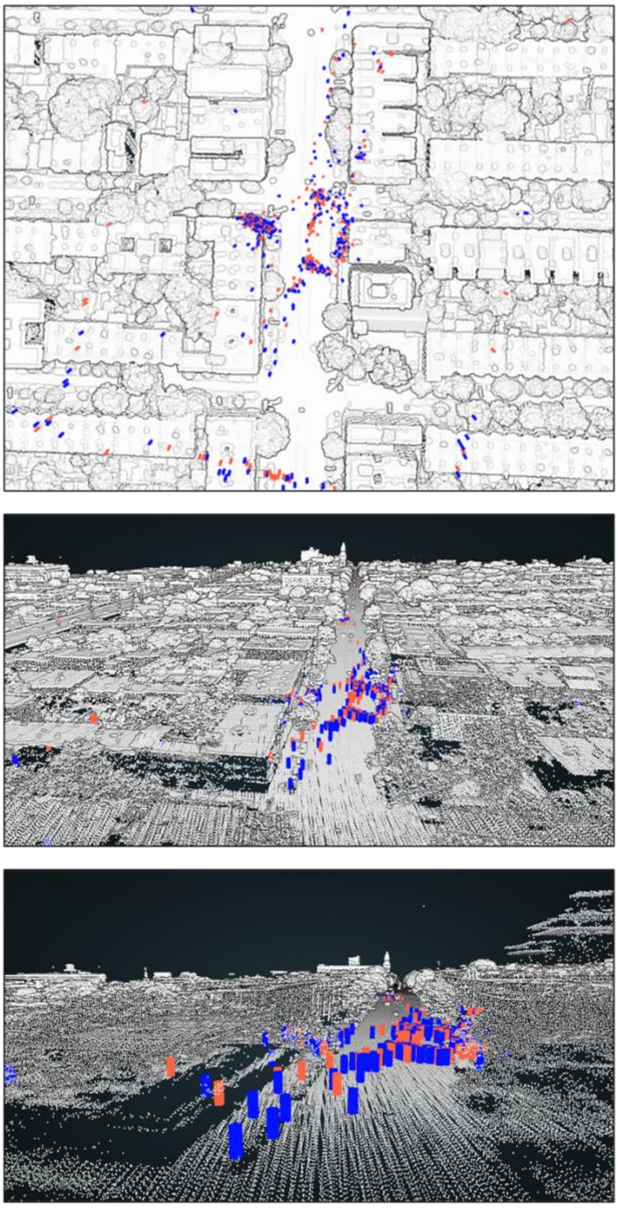


Figure 6. A set of 3D view of individuals in 3D space for the NYU Langone Hospital, Brooklyn, USA. Females in red and males in blue (Displayed data for all 1,578 points)

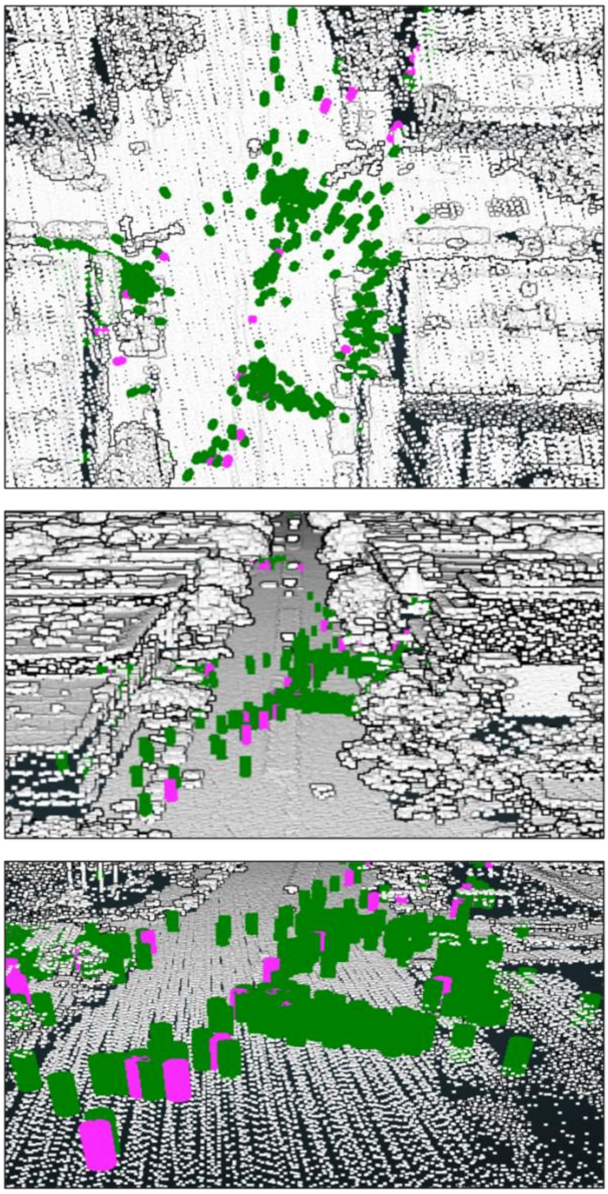


Figure 7. A set of 3D views of individuals in 3D space for NYU Langone Hospital, Brooklyn, USA. Medical personnel in green and others in pink (Displayed data for all 1578 points)

This research extended the prior work on 2D heatmap visualization proposed in Carey et al. (2021) through defining colors for every feature (e.g. time type, gender, medical Personnel, PPE usage) and specific color shading for attribute values within every feature. This was achieved using dynamic vertex color shading and through drop-down filters to select features of choice in the Potree dashboard. Color scale legend in the Potree dashboard updates automatically when a particular feature from the dropdown menu is selected.

This research proposed a novel methodology to visualize the data in 3D space by annotating the individuals as cylindrical polygons instead of 2D point maps. The concentrated co-existence of multiple individuals in specific locations around a healthcare facility depicted the high-risk areas for spread of the COVID-19 virus. These locations were labeled as interim and final destinations of observed individuals (e.g. coffee shop, subway stations) to highlight high risk points of interest. The 3D scene visualizations provided the realistic depth to understand the individuals' location and also allowed the viewers to see into spaces by zooming and moving around the spaces. For example, (1) the gender feature was displayed using different colors to distinguish between the movement patterns among males and females (Fig 6). (2) the medical personnel feature was selected, and the implemented color shading represented the differences in proportion of healthcare workers from non-medical personnel (Fig 7). The high volume of non-medical individuals when compared with the limited number of healthcare workers illustrated the dire situation during the COVID-19 pandemic with all the healthcare workers strained for capacity.

5. DISCUSSION

Human behaviors can help determine outbreak trajectories of infectious diseases such as COVID-19. This fundamental relationship underlies why behavioral interventions near high-risk zones (i.e. health care facilities) are effective tools in outbreak management. To understand the human behavior patterns and disease transmissibility, this research visualized the most affected areas at the time of COVID-19. Apart from healthcare facilities, coffee shops, public transport stands, parking lots were observed as high interaction areas where individuals congregated. To visualize the 3D point cloud data, point cloud tiles were rendered in Potree post-conversion in binary octree structure for managing the computation power. Millions of points in every point cloud tile put heavy load on the graphics memory of the system. Further application development could help optimize performance of Potree rendering by only loading necessary data in memory. For improving the visualizations, point cloud tiles can be sliced and merged with higher density, terrestrial scans, which may provide deeper insights and better system performance. During analysis, some healthcare facilities were found to be located on the edge of multiple point cloud tiles. To visualize these locations and understand individual behaviors, multiple files needed to be merged. Enhancing the visualization with additional information through text labels and feature overlays could help researchers further understand the data. As an alternative, the pathways could be treated as lines, instead of polygons, but that would lend itself to different exploration capabilities.

NSF DETER data are quite detailed (Laefer et al., 2020) and provide information on a per person basis (Laefer et al., 2021). This dataset provides a lot of scope to extend the research work in various directions. The implementation plan can be extended to 3D visualization of individuals to get a clear glimpse of behaviors across the streets. This can lead to better understanding of an individual's touch patterns while coming out of a facility by analyzing behaviors such as a touch on railing or a doorknob. While extending the work in future to comparative analysis, this data can help to show the most vulnerable businesses across the health care facility streets. This epidemic has had a major impact on the lifestyle of people and by understanding the epidemiological patterns, future disasters can be prevented through ahead of time planning. This paper lays out the initial footsteps for 3D visualization of individuals in ariel LiDAR scene setting. The next step in this study would be to show the areas of actual touch (e.g. handrails, parking meters, bus stand poles, trash cans etc.) by individual's around health care facilities. The research team will partner with colleagues in NYU's Global Health Program to help translate such capabilities and affiliated insights into practice.

6. CONCLUSIONS

This paper presents a means to automatically clean, classify and enable citizen science data and to contextualize the data through its coupling with aerial laser scans in an interactive platform where users can investigate the gathering and trajectories of specific groups of subjects (e.g. men versus women). Such enablement also helps visualize potentially, high-risk areas around the healthcare facilities where individuals egressing COVID-19 laden facilities tended to congregate. In this paper, 2D behavioral data is visualized in 3D point cloud space in 3D format by creating the cylindrical polygons to represent the individuals. The egress behavior of individuals in spatial context can be used to develop patterns for the spread of epidemic. Comparative analysis of visualizations of high-risk areas (COVID-19 facilities) can be used for the disease epidemiology and spread in the population. The research focused on bivariate features for analyzing egress behavior, however future work could incorporate gradient based color shading for multivariate features.

7. ACKNOWLEDGEMENTS

This work was supported by Bluefield GIS, the National Science Foundation (award #2027293), and NYU's Data Science and Software Services (DS3), which is funded by the Moore and Sloane foundations through the NYU Moore Sloane Data Science Environment.

8. REFERENCES

Adivar, B. and Selin Selen, E., 2013. Review of research studies on population specific epidemic disasters. *Disaster Prevention and Management: An International Journal*, 22(3), pp.243-264.

Buytaert, W., Dewulf, A., De Bièvre, B., Clark, J. and Hannah, D.M., 2016. Citizen science for water resources management: toward polycentric monitoring and governance?. *Journal of Water Resources Planning and Management*, 142(4), p.01816002.

Carey, C., Romero, J. and Laefer, D.F., 2021. New Potree shader capabilities for 3d visualization of behaviors near covid-19 rich healthcare facilities. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLVI-4/W4-2021, pp.61-66.

Console, R. and Carluccio, R., 2021. Earthquake Simulators Development and Application. *Statistical Methods and Modeling of Seismogenesis*, pp.27-62.

Danchilla, B., 2012. Three.js Framework. *Beginning WebGL for HTML5*, pp.173-203 (Accessed: 06/07/2022).

Danon, L., Brooks-Pollock, E., Bailey, M. and Keeling, M. (2021). A spatial model of COVID-19 transmission in England and Wales: early spread, peak timing and the impact of seasonality. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 376(1829). doi:10.1098/rstb.2020.0272.

Data.cityofnewyork.us. 2022. [online] Available at: <https://data.cityofnewyork.us/City-Government/Topobathymetric-LiDAR-Data-2017-/7sc8-jtbz> (Accessed: 06/07/2022).

Dlamini, W., Dlamini, S., Mabaso, S. and Simelane, S., 2020. Spatial risk assessment of an emerging pandemic under data scarcity: A case of COVID-19 in Eswatini. *Applied Geography*, 125, p.102358.

Kiatpanont, R., Tanlamai, U. and Chongstitvatana, P., 2016. Extraction of actionable information from crowdsourced disaster data. *Journal of emergency management*, 14(6), pp.377-390.

Laefer, D., Kirchner, T., Cheong, D., Khan, A., Qiu, W., Tai, N., Truong, T. and Virk, M., 2022. Data Resource Profile: Egress Behavior from Select NYC COVID-19 Exposed Health Facilities March-May 2020. [online] *arXiv.org*. Available at: <https://arxiv.org/abs/2101.10079>

Leal-Neto, O., Santos, F., Lee, J., Albuquerque, J. and Souza, W., 2020. Prioritizing COVID-19 tests based on participatory surveillance and spatial scanning. *International Journal of Medical Informatics*, 143, p.104263.

Lee, S. and Laefer, D., 2021. Spring 2020 COVID-19 community transmission behaviours around New York City medical facilities. *Infection Prevention in Practice*, 3(3), p.100158.

Li, W. and Rodriguez, P., 2008. Assessing the Impact of Hurricanes Katrina and Rita. *SSRN Electronic Journal*.

Melin, P., Monica, J.C., Sanchez, D. and Castillo, O. (2020). Analysis of Spatial Spread Relationships of Coronavirus (COVID-19) Pandemic in the World using Self Organizing Maps. *Chaos, Solitons & Fractals*, 138, p.109917. doi:10.1016/j.chaos.2020.109917.

Mohanty, S., Fairburn, S., Imhof, B., Ransom, S. and Vogler, A., 2008. Survey of Past, Present and Planned Human Space Mission Simulators. *SAE Technical Paper Series*.

Nielsen, C., Oberle, A. and Sugumaran, R., 2011. Implementing a High School Level Geospatial Technologies and Spatial Thinking Course. *Journal of Geography*, 110(2), pp.60-69.

OECD Regions and Cities at a Glance 2020, 2020. The changing shape of cities: Density and suburbanisation.

Paez, A., Lopez, F., Menezes, T., Cavalcanti, R. and Pitta, M., 2020. A Spatio-Temporal Analysis of the Environmental Correlates of COVID-19 Incidence in Spain. *Geographical Analysis*, 53(3), pp.397-421.

Palen, L. and Hughes, A.L., 2018. Social media in disaster communication. *Handbook of disaster research*, pp.497-518. Vancouver.

Potree.github.io. 2022. Potree. [online] Available at: <https://potree.github.io/> (Accessed: 06/07/2022).

Qgis.org. 2022. Welcome to the QGIS project! [online] Available at: <https://qgis.org/en/site/> (Accessed: 06/07/2022).

Reges, H.W., Doesken, N., Turner, J., Newman, N., Bergantino, A. and Schwalbe, Z., 2016. CoCoRaHS: The evolution and accomplishments of a volunteer rain gauge network. *Bulletin of the American Meteorological Society*, 97(10), pp.1831-1846.

St. John, M., Cowen, M., Smallman, H. and Oonk, H., 2001. The Use of 2D and 3D Displays for Shape-Understanding versus Relative-Position Tasks. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 43(1), pp.79-98.

Usgs.gov. 2022. 3D Elevation Program | U.S. Geological Survey. [online] Available at: <https://www.usgs.gov/3d-elevation-program> (Accessed: 06/07/2022).

Xie, Z., Qin, Y., Li, Y., Shen, W., Zheng, Z. and Liu, S., 2020. Spatial and temporal differentiation of COVID-19 epidemic spread in mainland China and its influencing factors. *Science of The Total Environment*, 744, p.140929.

Zhao, R., Pang, M. and Wang, J., 2018. Classifying airborne LiDAR point clouds via deep features learned by a multi-scale convolutional neural network. *International Journal of Geographical Information Science*, 32(5), pp.960-979.