

Utilizing Mobility Tracking to Identify Hotspots for Contagious Disease Spread

A Case Study of UNITEN Students Using Google Map Data

Yaw Mei Wyin

Institute of Sustainable Energy, Universiti Tenaga Nasional, 43000
Kajang, Malaysia

Prajindra Sankar Krishnan

Institute of Sustainable Energy, Universiti Tenaga Nasional, 43000
Kajang, Malaysia

Chen Chai Phing

Institute of Sustainable Energy, Universiti Tenaga Nasional, 43000
Kajang, Malaysia

Tiong Sieh Kiong

Institute of Sustainable Energy, Universiti Tenaga Nasional, 43000
Kajang, Malaysia

Abstract: A significant global health problem nowadays is the incidence of serious infectious illnesses. An extraordinary humanitarian crisis has been brought on by the current COVID-19 pandemic, which has spread around the world. The spread of new viruses has put established healthcare institutions under tremendous strain and created a number of pressing problems. It is important to predict the future movement and pattern of the illness in order to decrease infectious instances and maximize recovered cases. This research paper aims to utilize mobility tracking as a means to identify hotspots for contagious disease spread. The study focuses on collecting and analyzing mobility data from UNITEN students using Google Map data over a period of two weeks. The paper describes the data collection process, data pre-processing steps, and the application of the HDBSCAN algorithm for hotspot clustering. The results demonstrate the effectiveness of HDBSCAN in identifying hotspots based on the mobility data. The findings highlight the potential of mobility tracking for disease surveillance and provide insights for public health interventions and preventive measures.

Keywords: Hotspot, HDBSCAN, infection disease, COVID-19.

Introduction

In particular for highly contagious illnesses like COVID-19, human mobility is a crucial consideration in the understanding of an infectious disease's transmission ([Gatto et al., 2020](#); [Kiang et al., 2021](#); [Venkatramanan et al., 2021](#); [Wang et al., 2020](#); [Wu, Leung & Leung, 2020](#)). Early patient identification and the implementation of prophylactic measures are the most effective ways to reduce the disease's rate of transmission during an outbreak. It is now feasible to follow people's movements over extended periods of time and across a large region because of the proliferation of gadgets with accurate localization capabilities and robust wireless networks.

Humans can contract infectious illnesses either directly or indirectly through time and distance. For instance, COVID-19 is regarded as illnesses that may be conveyed to vulnerable individuals directly when airborne droplets from an infected person's coughing or sneezing travel a short distance ([Setti et al., 2020](#); [Devi et al., 2022](#); [Zarei et al., 2021](#)). As a result, effective transmissions between people require a close enough physical distance.

When contemplating the treatment of infectious illnesses, spatial dimensions are essential. An outbreak of an epidemic is defined from an epidemiological perspective as any temporally anomalous rise in the number of case-patients in a specific location ([Porta, 2014](#); [CDC, 2018](#)). Effective disease management and prevention techniques depend on pinpointing areas where infectious diseases are most likely to spread. Contagious diseases pose significant threats to public health, requiring proactive strategies for containment and prevention. One crucial aspect of disease control is the identification of hotspots, geographic areas with a high concentration of disease transmission. Geographical regions or localities known as hotspots are those where an infectious illness is spreading particularly rapidly or heavily.

Researchers now have access to a variety of databases that monitor people's movements, enabling more thorough analysis and epidemic spread models at a finer geographical and temporal resolution. GPS and Google Location History are examples of conventional sources of mobility data where several sorts of motions may be collected ([Ruktanonchai et al., 2018](#)).

This research paper explores the use of mobility tracking as a tool for hotspot identification. Mobility data, collected from UNITEN (Universiti Tenaga Nasional) students through Google Maps, provides valuable insights into individuals' movement patterns and their potential impact on disease spread. By analyzing this data, we aim to identify hotspots and understand the role of mobility in contagious disease transmission.

The software platform Google History Location Extractor is suggested in this article. The platform simulates the epidemic spreading using real-world UNITEN student trajectory

datasets, estimates the contact behaviours using GPS trajectories, and simulates their mobility in response to public policies, to be able to assess how well they work to halt the epidemic's spread. The platform, which combines the probabilistic model of individual-level infectious disease transmission with the investigation of individual infection risks, also facilitates data research and mining on the risks of infection spreading, such as the identification of possible secondary contacts.

The next sections cover relevant research that has already been done and go into great depth on our movement tracking system as a tool for locating hotspots. The other sections of the study detail our findings and include the modelling of mobility patterns, the AI algorithm implementation procedure, outcomes, and results. We summarize the main points and emphasize the importance of our platform in our conclusion.

Related Works

Understanding disease transmission in a population may be done by building a network to reflect person-to-person interactions. Various techniques, including surveys, statistical methods, census data, and movement data ([Chang et al., 2021](#); [Kumar, Oke & Nahmias-Biran, 2021](#); [Maheshwari & Albert, 2020](#); [Muller et al., 2020](#); [Rechtin et al., 2020](#); [Schlosser et al., 2020](#); [Soures et al., 2020](#); [Yi et al., 2021](#)), can be used to create these networks. People's positions and the amount of time they spend in certain areas are detailed in a vast quantity of data as a result of the widespread usage of location tracking applications on mobile phones. In response to COVID-19, new systems like Google Community Mobility Reports ([Google, 2020](#)) provide devices and aggregated movement data. Data is compiled and made anonymous to safeguard people's privacy.

In order to emphasize geographical patterns of illnesses and determine whether there is a pattern of disease incidence in a certain region, spatial analyses and statistics, such as spatial autocorrelation analysis, cluster analysis, and temporal analysis, are frequently utilised ([Brownstein et al., 2002](#); [Tsai et al., 2009](#); [Pace, Barry & Sirmans, 1998](#); [Ping et al., 2004](#)). Growing interest in the identification of disease clusters or "hotspots" for public health surveillance, particularly to better understand the rising prevalence of dengue fever, has been sparked by recent developments in spatial statistics in geographical information systems ([Yeshiwondim et al., 2009](#)). The dynamics of disease propagation can be understood through the analysis of spatiotemporal patterns. So as to target disease monitoring and management in places and periods with increased risk for illness, it is helpful to detect geographical, temporal, and space-time clustering ([Si et al., 2008](#)).

Previous research has investigated various methods for hotspot identification and their applications in disease spread analysis. Studies have utilized data from different sources, such

as mobile phones, GPS tracking, and social media platforms, to understand human mobility patterns and their association with disease transmission. However, there are still limitations and gaps in the current research, including the need for more accurate and granular data, effective clustering algorithms, and comprehensive analysis techniques.

Methodology

Epidemiologists can foresee disease outbreaks by comprehending the patterns of human movement, since infectious diseases are disseminated through direct human contact. Understanding how people move through time and space might help us respond to and recover from such crisis situations. This can be done with the use of mobile phone location data. One of the crucial components that defines these mobilities is location history. The cell site location method, GPS, and Wi-Fi positioning are some of the methods ([Khalel, 2010](#)) that may be used to locate someone.

All three of these technologies are used by Google Maps to locate the user's position with great accuracy. Users may export this data via the Google Takeout service, which provides the data in JSON (JavaScript Object Notation) format. The produced JSON file contains a wide variety of different kinds of metadata. The direction, activity type, latitudeE7 (latitude), longitudeE7 (longitude), precision (accuracy measure), timestampMS (timestamp in milliseconds), and altitude are some of the most crucial data items that are highlighted.

Google Maps has several capabilities in addition to place searches and route finding. Google Timeline, a function of Google Maps, constantly tracks user behaviour, including the places visited and the method of transportation used to get there, using GPS position data. This raw data reveals the user's location at a certain time and date.

The methodology employed in this study involved the utilization of mobility tracking and Google Maps data to identify hotspots for contagious disease spread among UNITEN students. At the beginning of the study, mobility data from 50 UNITEN students were collected using the Google Maps application in the mobile phone over two-week periods, as illustrated in Figure 1.

MovementMode	UserID	startLocation_Lat	startLocation_Long	endLocation_Lat	endLocation_Long	duration_Start	duration_End	distance	activityType	confidence
ActivitySegment	20230614163218	3.708416	103.3271788	3.6927028	103.3390561	2022-01-06 07:4...	2022-01-06 07:4...	2188	MOTORCYCLING	LOW
ActivitySegment	20230614163218	3.708416	103.3271788	3.8205583	103.3256068	2022-01-06 08:2...	2022-01-06 09:0...	12470	IN_PASSENGER...	HIGH
ActivitySegment	20230614163218	3.8202583	103.3256371	3.8211395	103.3184758	2022-01-06 09:5...	2022-01-06 10:0...	800	IN_PASSENGER...	HIGH
ActivitySegment	20230614163218	3.8198361	103.3231609	3.8342668	103.3024717	2022-01-09 14:3...	2022-01-09 14:5...	2800	IN_PASSENGER...	LOW
PlaceVisit	20230614163218	3.8343991	103.3023184			2022-01-09 14:5...	2022-01-09 16:1...	0	Sizzling Station G...	84.4508
ActivitySegment	20230614163218	3.8342879	103.3020501	3.8216177	103.3181615	2022-01-09 16:1...	2022-01-09 16:2...	2275	IN_PASSENGER...	HIGH
PlaceVisit	20230614163218	3.8206798	103.3178493			2022-01-09 16:2...	2022-01-09 16:4...	0		56.785046
PlaceVisit	20230614163218	3.8206798	103.3178493			2022-01-13 08:4...	2022-01-14 06:2...	0		48.003323
ActivitySegment	20230614163218	3.8211596	103.3184932	3.8171501	103.3302818	2022-01-14 06:2...	2022-01-14 08:3...	4009	IN_PASSENGER...	MEDIUM
PlaceVisit	20230614163218	3.817229	103.330138			2022-01-14 08:3...	2022-01-14 08:4...	0	Gurtefing® Barber...	32.1175
ActivitySegment	20230614163218	3.8163351	103.3299244	3.8256222	103.3282423	2022-01-14 08:4...	2022-01-14 08:4...	1243	IN_PASSENGER...	HIGH
PlaceVisit	20230614163218	3.8206798	103.3178493			2022-01-15 15:3...	2022-01-15 17:3...	0		49.35097
PlaceVisit	20230614163218	3.8276312	103.3259872			2022-01-25 10:0...	2022-01-25 13:2...	0	Eco Save RM2	18.229351
ActivitySegment	20230614163218	3.8268644	103.3258479	3.8211734	103.3185306	2022-01-25 13:2...	2022-01-25 13:3...	1029	IN_PASSENGER...	LOW
PlaceVisit	20230614163218	3.8206798	103.3178493			2022-01-25 13:3...	2022-01-25 19:0...	0		56.374504
PlaceVisit	20230614163218	3.8259201	103.303955			2022-01-27 15:1...	2022-01-27 15:2...	0	McDonald's Inder...	92.1545
ActivitySegment	20230614163218	3.8255225	103.3045683	3.8198382	103.3259288	2022-01-27 15:2...	2022-01-27 16:0...	2452	IN_PASSENGER...	MEDIUM
PlaceVisit	20230614163218	3.8195488	103.3259658			2022-01-27 16:0...	2022-01-27 16:2...	0	The Zenith Hotel	79.175964
PlaceVisit	20230614163218	3.8402509	103.3318991			2022-01-29 04:0...	2022-01-29 04:1...	0	Semanambu Badmi...	90.89959
ActivitySegment	20230614163218	3.8395644	103.3312121	3.8182077	103.3071673	2022-01-29 04:1...	2022-01-29 04:2...	3571	IN_PASSENGER...	HIGH
PlaceVisit	20230614163218	3.8178277	103.3065581			2022-01-29 04:2...	2022-01-29 04:3...	0	HoHo Hainan Ko...	49.521767

Figure 1. Extracted Data Display

In order to identify potential COVID-19 dissemination, the location records were examined using the Google History Location Extractor and Indicator. The study involves plotting the locations of the students on a map with complete information about the date, time, region, and distance between persons with the aim of looking for prospective COVID-19 distribution scenarios. Figure 2 displays the Venn chart that was used to analyse the location information for two students.

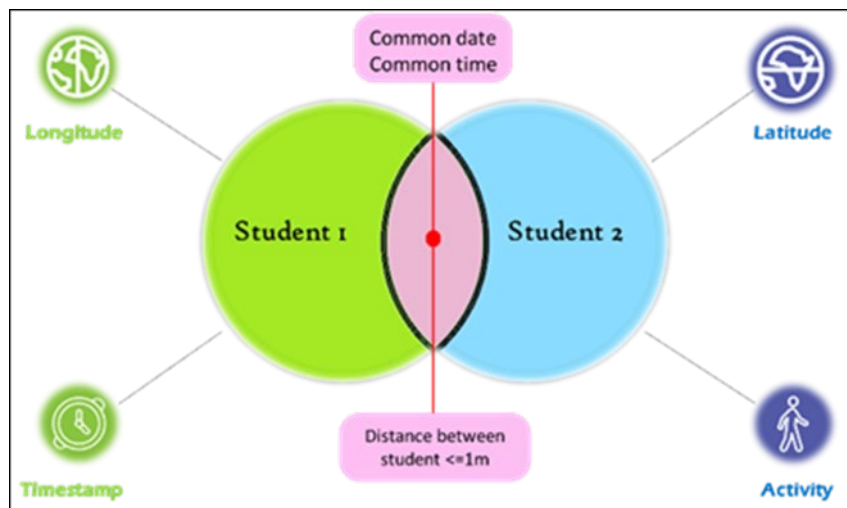


Figure 2. Venn diagram of location analysis for two students

This study looked at the JSON file containing the location history information of 50 students using tools to extract and visualise the downloaded data. The Google History Location Extractor program, which is depicted in Figure 3, performs a variety of tasks, including data extraction, graphical user interface presentation, data saving in CSV or KML format, and mobility visualisation on Google Maps. The data may also be exported to a Microsoft SQL Server database format for additional in-depth analysis. The full software program’s flowchart,

which includes data collecting, import, filtering, distance calculation, and visualization, is shown in Figure 4. The overall algorithm flow visualization is shown in Figure 5.

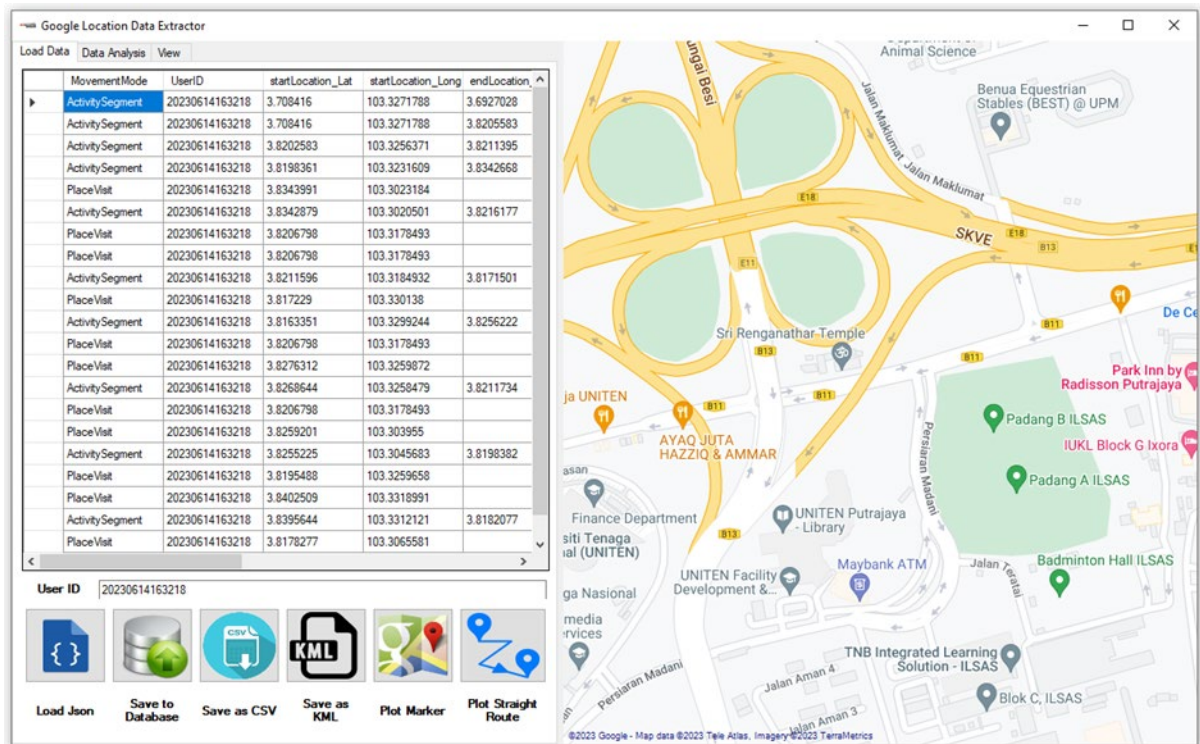


Figure 3. GUI of Google Historical Location Extractor software

Since it was simpler to manage using Google Timeline, the location history data was collected in JSON rather than KML (Sardianos, Varlamis & Bouras, 2018). Through the use of the import method in VB.net, the JSON file is transformed into a table format. The needed variable may be chosen, variable names can be changed, and the data type for variables can be specified while the data is still being imported. Latitude, longitude, accuracy, and timestamp are filtered from the raw data for this project and kept separately in tabular format. The mobility data was loaded from the 'MobilityData_49.csv' file into a Pandas DataFrame. To focus on a specific area of interest, the data was spatially filtered based on bounding box coordinates. The longitude and latitude ranges were defined to select the relevant data points. Following the filtering process, the data underwent pre-processing steps. The 'duration_Start' and 'duration_End' columns were converted to datetime format using the `pd.to_datetime()` function, enabling the calculation of stay durations. The stay duration for each data point was determined by computing the difference between the 'duration_End' and 'duration_Start' timestamps and converting the result to minutes.

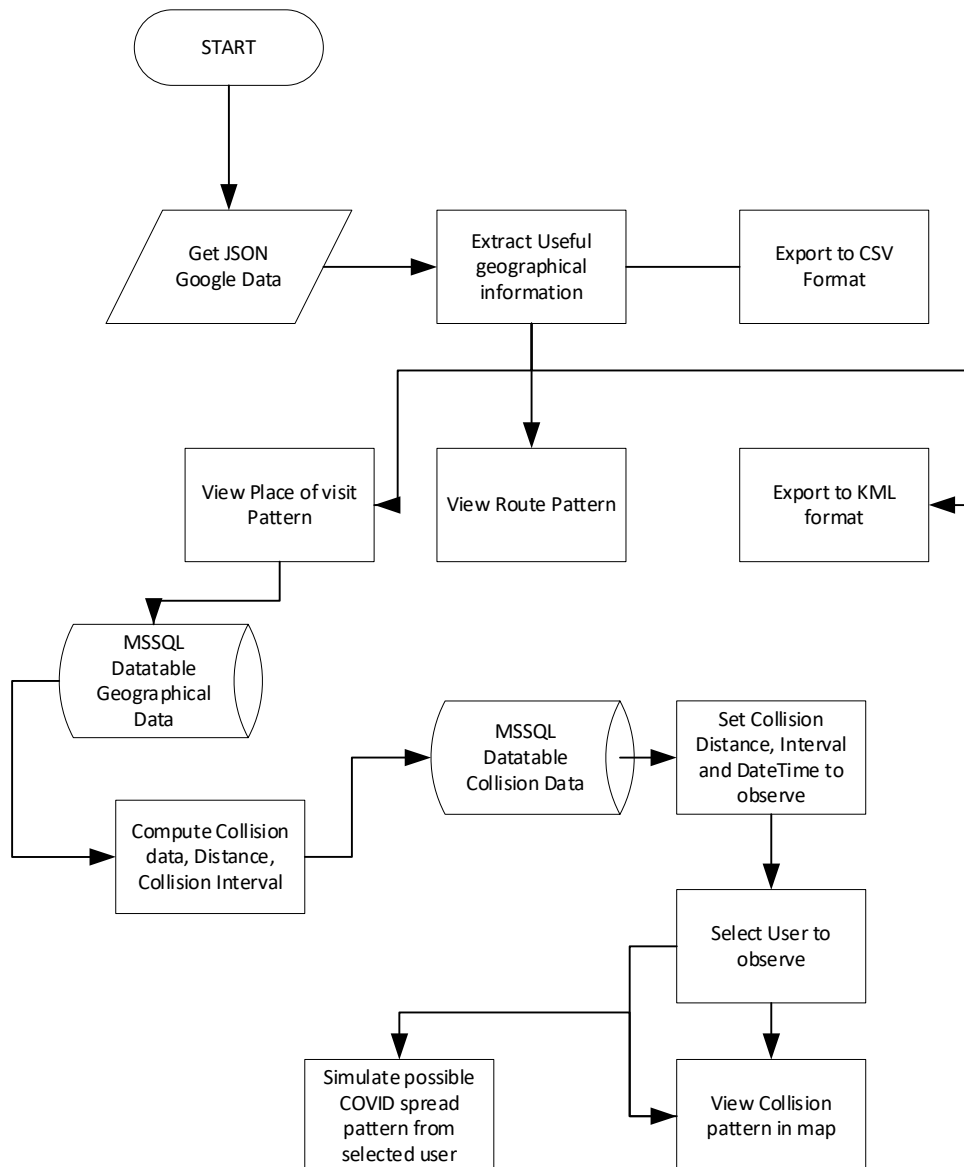


Figure 4. Entire system architecture

This study employs the HDBSCAN (Hierarchical Density-Based Spatial Clustering of Applications with Noise) algorithm to identify hotspots for contagious disease spread using mobility tracking data. HDBSCAN is an extension of the DBSCAN (Density-Based Spatial Clustering of Applications with Noise) algorithm, which is known for its ability to discover clusters of arbitrary shapes and handle noise effectively.

The HDBSCAN algorithm operates by defining clusters based on the density of data points. It starts by calculating the mutual reachability distance between each pair of points, which combines both the distance between points and their local densities. This distance measure allows the algorithm to identify regions of higher density, which are likely to represent meaningful clusters.

In the context of this research, the data points are denoted as $X = x_1, x_2, \dots, x_n$, where x_i represents the coordinates of the i -th data point.

- Density Calculation:**
 The density of a data point x_i is calculated as follows:

$$D(x_i) = \sum_{j=1}^n \delta(\text{dist}(x_i, x_j), \epsilon),$$
 where δ is the Kronecker delta function and $\text{dist}(x_i, x_j)$ represents the distance between data points x_i and x_j . ϵ is the specified radius within which neighboring points are considered.
- Mutual Reachability Distance:**
 The mutual reachability distance between points x_i and x_j is computed as follows:

$$MRD(x_i, x_j) = \max(RD(x_i, x_j), RD(x_j, x_i)),$$
 where $RD(x_i, x_j)$ represents the reachability distance from x_i to x_j , and $RD(x_j, x_i)$ represents the reachability distance from x_j to x_i .
- Reachability Distance:**
 The reachability distance between points x_i and x_j is given by:

$$RD(x_i, x_j) = \max(D(x_j), \text{dist}(x_i, x_j)),$$
 where $D(x_j)$ represents the density of data point x_j , and $\text{dist}(x_i, x_j)$ represents the distance between x_i and x_j .
- Constructing the Condensed Tree:**
 The condensed tree is constructed by merging clusters based on their mutual reachability distances. Various techniques, such as minimum spanning tree or single linkage clustering, can be used to create the tree structure.
- Cluster Extraction:**
 The final set of clusters is extracted from the hierarchy determined by the condensed tree. The appropriate level of density threshold is determined to define clusters, considering the minimum cluster size parameter.

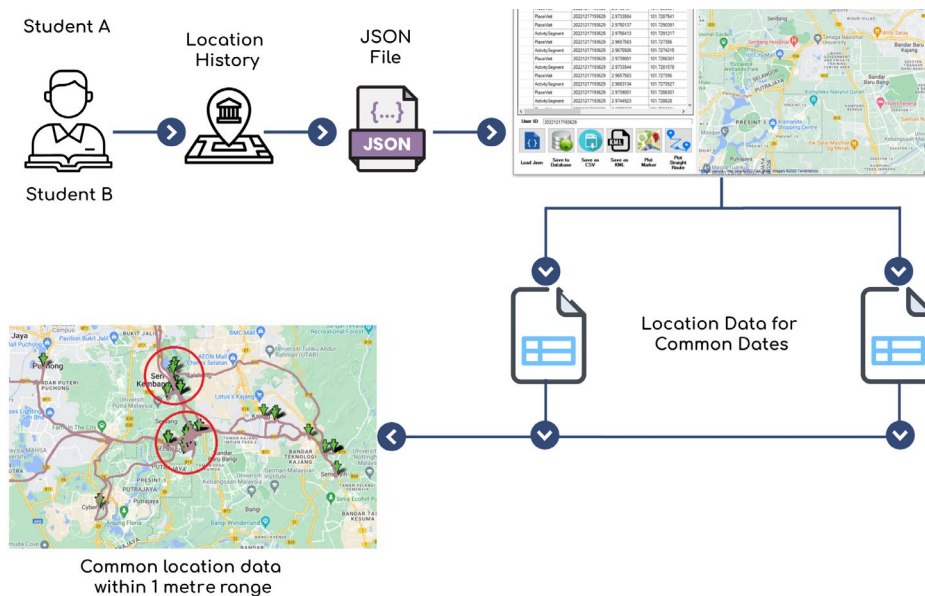


Figure 5. General flow of algorithm

In this research, mobility tracking data from Google Maps is first collected and filtered based on predefined bounding box coordinates to focus on the specific area of interest, the UNITEN campus as presented in Figure 6. The data is then pre-processed by converting the timestamps

to datetime format and calculating the duration of stay at each location. These steps enable the data to be effectively utilized by the HDBSCAN algorithm.

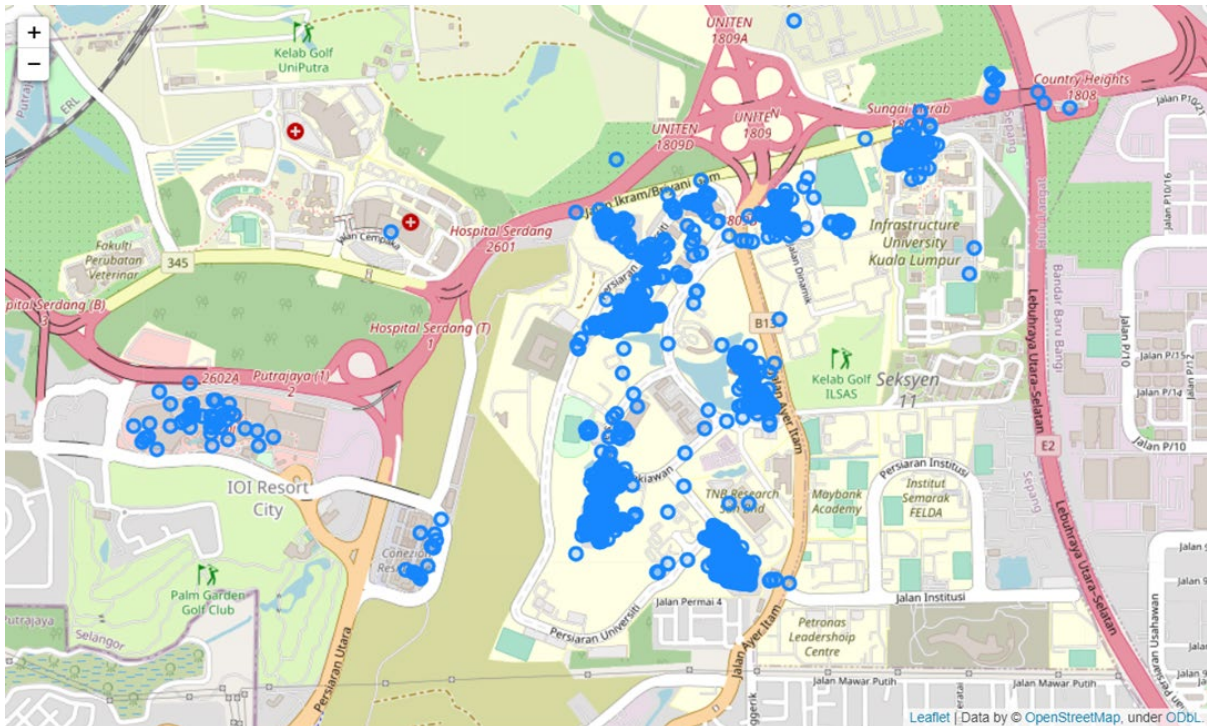


Figure 6. Students' GPS for area of interest after applying predefined bounding box

Next, the HDBSCAN algorithm is applied to the pre-processed mobility data, considering the coordinates of each data point. The algorithm takes into account the minimum cluster size parameter, which determines the minimum number of points required to form a cluster. This parameter ensures that only significant clusters, composed of a sufficient number of points, are identified as hotspots. The algorithm assigns each data point to a cluster or labels it as noise if it does not meet the density requirements.

HDBSCAN algorithm was used for hotspot identification because it is well-suited for clustering spatial data and requires the specification of parameters such as the minimum cluster size. The coordinates, represented by the 'startLocation_Lat' and 'startLocation_Long' columns, were used as input for the clustering algorithm. The resulting cluster labels were added as a new column ('Cluster') to the filtered data, facilitating the identification of hotspots.

After the clustering process, the identified clusters are analysed to determine the total number of hotspots within the UNITEN campus. The distribution of these hotspots across the study area is visualized using maps and charts. Additionally, the average time spent in each hotspot by all users is calculated and discussed, providing insights into the temporal dynamics of hotspot activity.

Results and Discussion

The analysis of the mobility data and hotspot identification yielded several key findings. A total of 18 hotspots were identified within the study area, signifying areas with a high concentration of potential disease transmission, as shown in Figure 7. These hotspots provide valuable insights into the locations that require particular attention in terms of disease surveillance and control efforts.

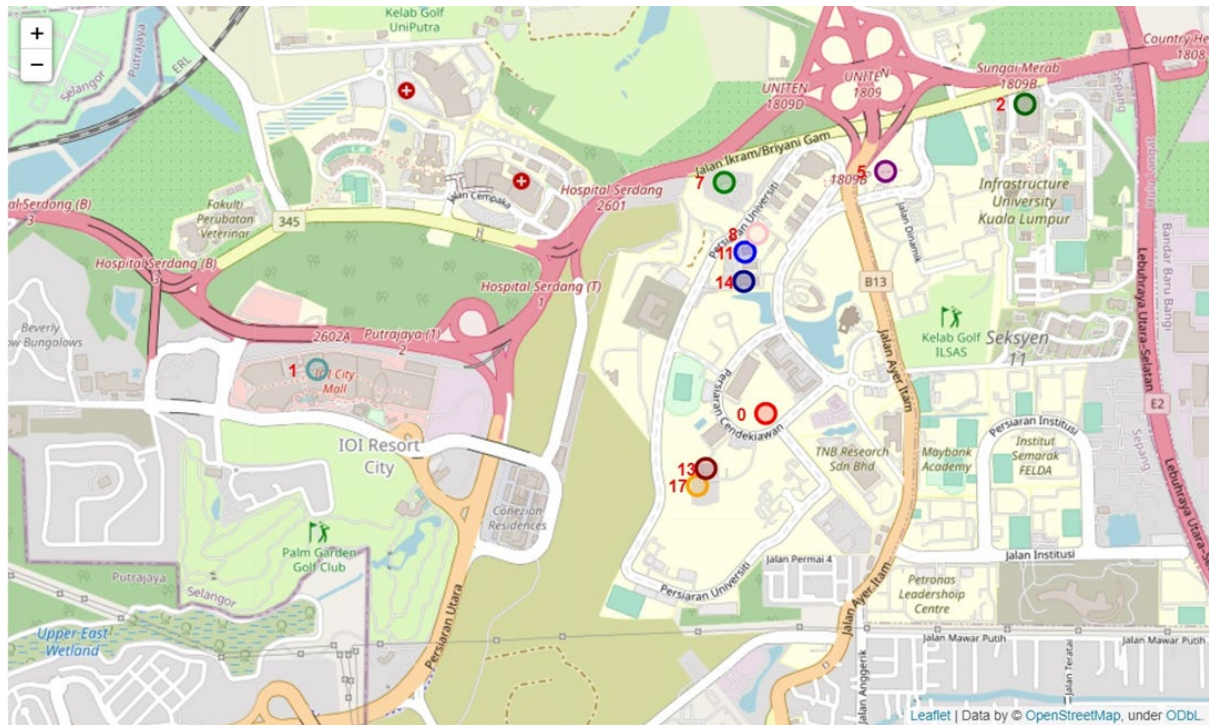


Figure 7. Hotspot identification

The average time spent in each hotspot by all users was also computed in Table 1. Figures 8 and 9 illustrate the sample of possible student collisions within the same day and same hour. This metric offers an understanding of the relative significance of specific locations in terms of duration. By identifying hotspots with longer average stay durations, public health authorities can prioritize these areas for targeted interventions and preventive measures.

Table 1. Average time spent (minutes) in each hotspot by all users for each cluster

Cluster	Average time spent in each hotspot by all users (minutes)
0	157.25
1	30.02
2	22.46
3	39.79
4	53.28
5	77.51
6	388.41
7	33.70
8	292.77
9	209.63

Cluster	Average time spent in each hotspot by all users (minutes)
10	106.73
11	68.97
12	352.74
13	115.84
14	121.03
15	546.49
16	497.05
17	575.08

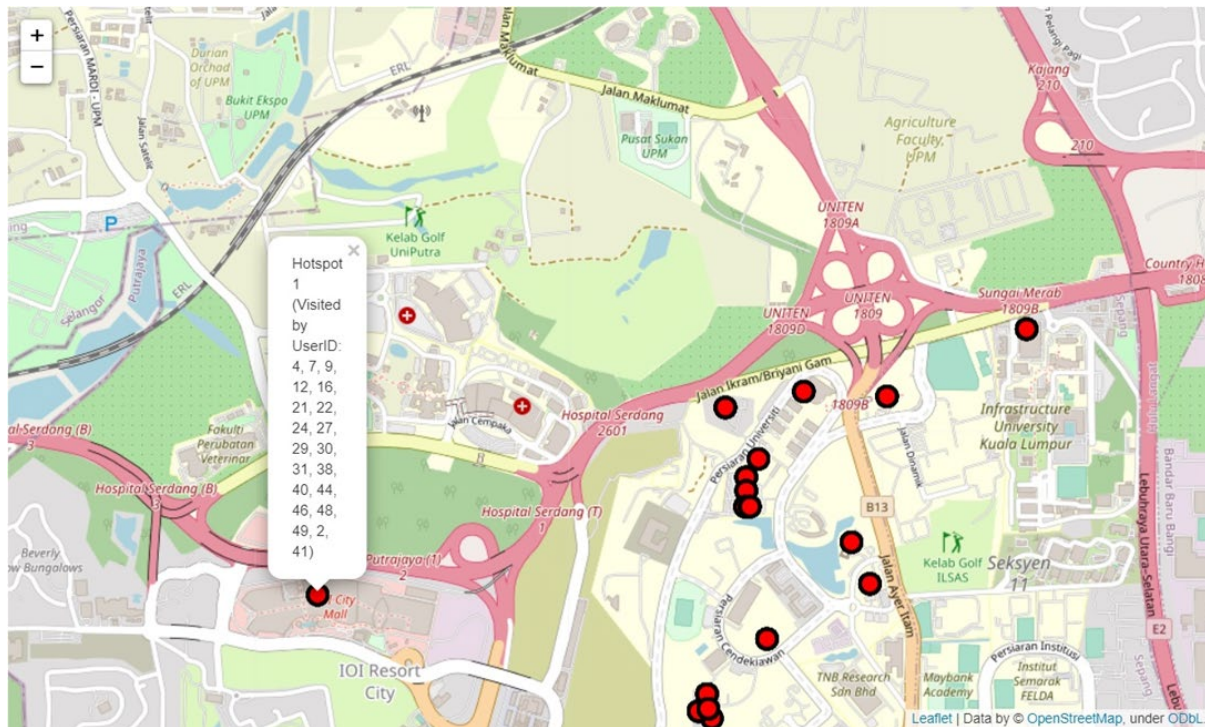


Figure 8. Sample of possible student collisions within the same day and same hour for hotspot 1

To visually represent the identified hotspots, a map was generated using the Folium library. Each hotspot was marked with a distinct colour and labelled with the corresponding hotspot number. This visualization allows for a clear understanding of the spatial distribution of hotspots and their relative proximity to each other.

Furthermore, the analysis focused on a specific user, UserID 1, and filtered the data for their visits from Monday to Sunday. By examining the unique hotspots visited by UserID 1, it was possible to identify the specific locations that this individual frequented. These visited hotspots were marked on the map using black circles filled in red, providing insights into the movement patterns and potential exposure risks for UserID 1, as illustrated in Figure 10. The same analysis was done for other UserIDs. Figure 11 is a sample of hotspots visited by UserID 2.

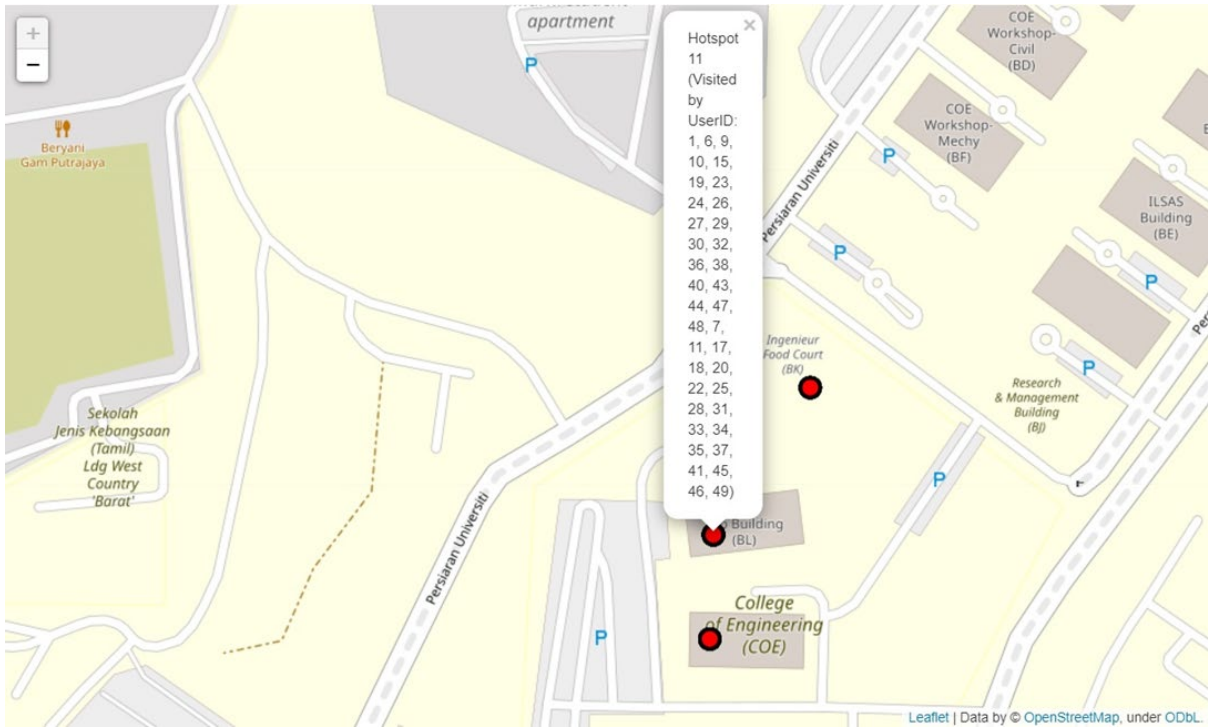


Figure 9. Sample of possible student collision within same day and same hour for hotspot 11

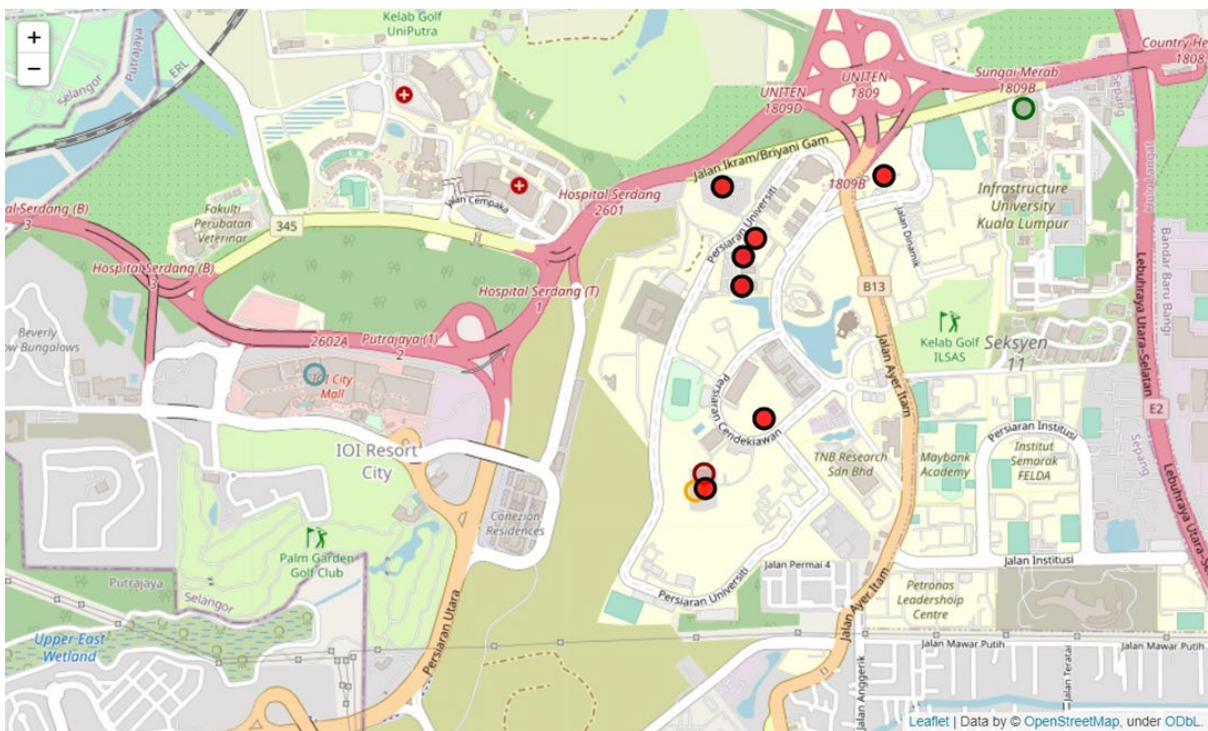


Figure 10. Hotspot visited by UserID 1

The combination of hotspot identification, average stay durations, and user-specific analysis offers a comprehensive understanding of the contagious disease spread patterns among UNITEN students. This information can inform public health strategies, interventions, and preventive measures tailored to mitigate the risk of disease transmission within the university community.

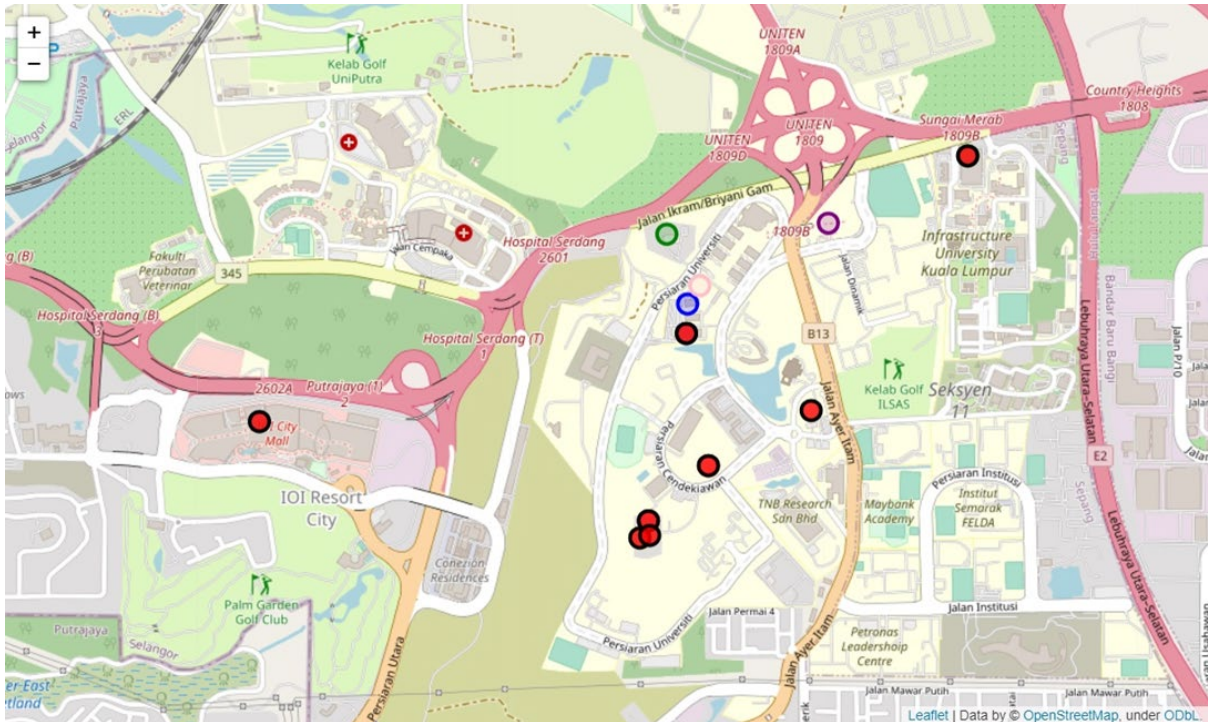


Figure 11. Hotspot visited by UserID 2

By applying the HDBSCAN algorithm to the mobility tracking data, this study leverages its ability to capture clusters of varying densities and irregular shapes. This allows for the identification of hotspots, representing areas of concentrated activity, which are crucial for understanding the potential spread of contagious diseases. Coordinates of HDBSCAN clustering with and without outliers are depicted in Figures 12 and 13.

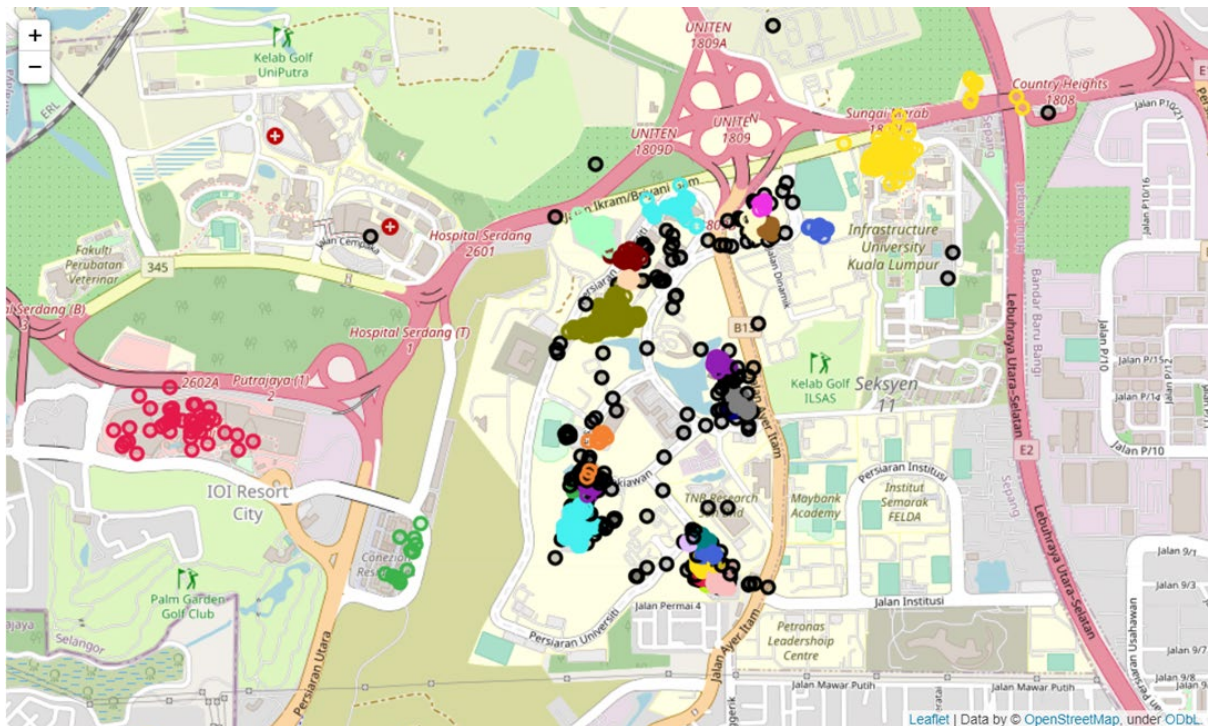


Figure 12. HDBSCAN clustering with outliers

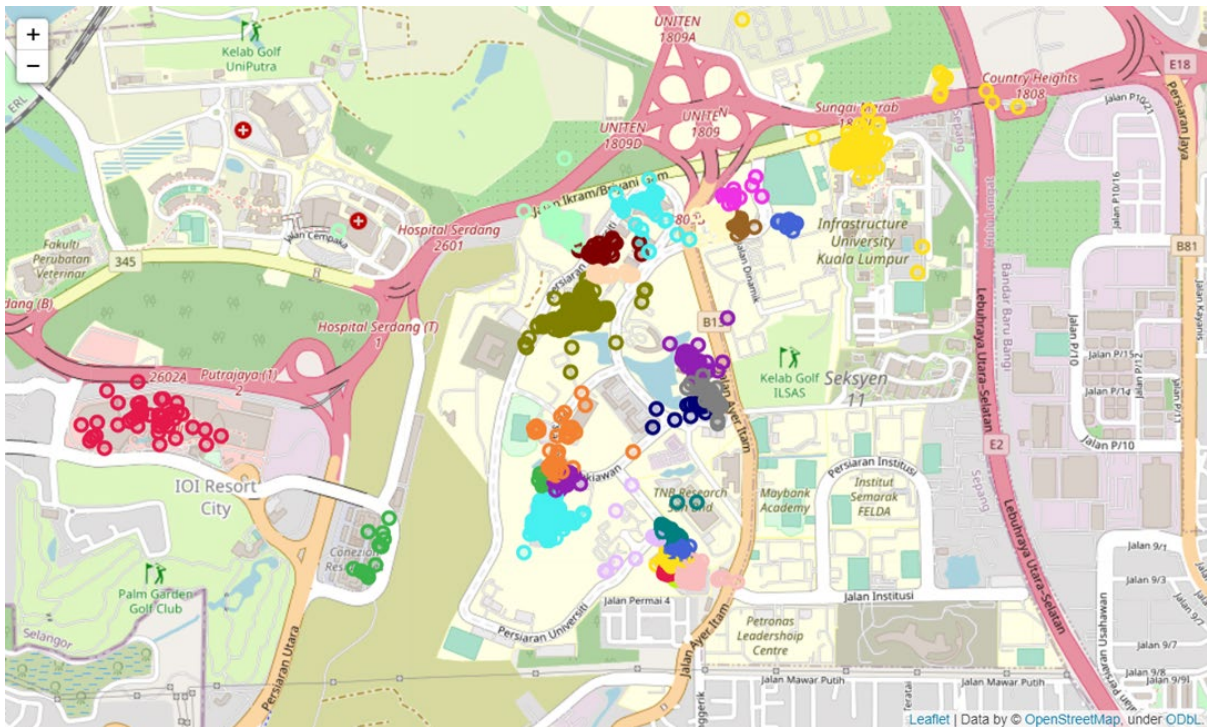


Figure 13. HDBSCAN clustering after removing outliers

In summary, the HDBSCAN algorithm is utilized in this research to effectively identify hotspots for contagious disease spread based on mobility tracking data. Its capability to handle density-based clustering and noise detection contributes to the accurate identification and characterization of hotspots within the UNITEN campus.

The findings of this research contribute to the understanding of contagious disease spread in relation to mobility patterns. The interpretation of the results in the context of disease transmission highlights the importance of hotspot identification for effective disease surveillance. The analysis of the effectiveness of mobility tracking and hotspot identification provides insights into the potential of these techniques for proactive public health interventions. The implications of the results include the development of targeted prevention strategies, resource allocation, and the implementation of timely interventions in hotspot areas.

Conclusion

This research has certain limitations that should be acknowledged. First, the data collection was limited to UNITEN students using Google Maps, which may not represent the entire population. Second, the algorithm used for hotspot clustering, HDBSCAN, has specific parameters that might affect the results. Future research should consider incorporating additional data sources to improve the accuracy and granularity of hotspot identification. Exploring different clustering algorithms and validation techniques could also enhance the robustness of the hotspot identification process.

In conclusion, this research demonstrates the potential of mobility tracking to identify hotspots for contagious disease spread. The analysis of mobility data collected from UNITEN students using Google Maps provides valuable insights into hotspot distribution and the role of mobility in disease transmission. The findings contribute to the field of contagious disease spread analysis and hotspot identification, highlighting the significance of proactive surveillance and targeted interventions. This research has the potential to impact public health strategies by enabling more effective measures to prevent and control the spread of contagious diseases.

Acknowledgements

This work is supported by IDRC Grant No. 109586 – 001 for Artificial Intelligence Framework for Threat Assessment and Containment for COVID-19 and Future Epidemics while Mitigating the Socioeconomic Impact to Women, Children and Underprivileged Groups.

A version of this paper was presented at the third International Conference on Computer, Information Technology and Intelligent Computing, CITIC 2023, held in Malaysia on 26–28 July 2023.

References

- Brownstein, J. S., Rosen, H., Purdy, D., Miller, J. R., Merlino, M., Mostashari, F., & Fish, D. (2002). Spatial analysis of West Nile virus: rapid risk assessment of an introduced vector-borne zoonosis. *Vector-Borne and Zoonotic Diseases*, 2(3), 157–164. <https://doi.org/10.1089/15303660260613729>
- CDC [US Centers for Disease Control and Prevention]. (2018) Managing HIV and Hepatitis C Outbreaks among People Who Inject Drugs - A Guide for State and Local Health Departments. Available from <https://www.cdc.gov/hiv/pdf/programresources/guidance/cluster-outbreak/cdc-hiv-hcv-pwid-guide.pdf>
- Chang, S., Pierson, E., Koh, P. W., Gerardin, J., Redbird, B., Grusky, D., & Leskovec, J. (2021). Mobility network models of COVID-19 explain inequities and inform reopening. *Nature*, 589, 82–87. <https://doi.org/10.1038/s41586-020-2923-3>
- Devi, S., Nagaraja, K. V., Thanuja, L., Reddy, M. V., & Ramakrishna, S. (2022). Finite element analysis over transmission region of coronavirus in CFD analysis for the respiratory cough droplets, *Ain Shams Engineering Journal*, 13(6), 101766. <https://doi.org/10.1016/j.asej.2022.101766>
- Gatto, M., Bertuzzo, E., Mari, L., Miccoli, S., Carraro, L., Casagrandi, R., & Rinaldo, A. (2020). Spread and dynamics of the COVID-19 epidemic in Italy: Effects of emergency containment measures. *Proceedings of the National Academy of Sciences (USA)*, 117(19), 10484–10491. <https://doi.org/10.1073/pnas.2004978117>
- Google. (2020). COVID-19 Community Mobility Reports, 2020. Available from <https://www.google.com/covid19/mobility>

- Khalel, A. M. H. (2010). Position Location Techniques in Wireless Communication Systems (Dissertation). Retrieved from <http://urn.kb.se/resolve?urn=urn:nbn:se:bth-4796>
- Kumar, N., Oke, J., & Nahmias-Biran, Bh. (2021). Activity-based epidemic propagation and contact network scaling in auto-dependent metropolitan areas. *Scientific Reports*, 11, 22665. <https://doi.org/10.1038/s41598-021-01522-w>
- Maheshwari, P., & Albert, R. (2020). Network model and analysis of the spread of Covid-19 with social distancing. *Applied Network Science*, 5, 100. <https://doi.org/10.1007/s41109-020-00344-5>
- Kiang, M. V., Santillana, M., Chen, J.T., Onnela, J.-P., Krieger, N., Engø-Monsen, K., Ekapirat, N., Areechokchai, D., Prempre, P., Maude, R. J., & Buckee, C. O. (2021). Incorporating human mobility data improves forecasts of Dengue fever in Thailand. *Scientific Reports*, 11, 923. <https://doi.org/10.1038/s41598-020-79438-0>
- Muller, S. A., Balmer, M., Neumann, A., & Nagel, K. (2020). Mobility traces and spreading of COVID-19. medRxiv preprint. <https://doi.org/10.1101/2020.03.27.20045302>
- Pace, R. K., Barry, R., & Sirmans, C. F. (1998). Spatial Statistics and Real Estate. *The Journal of Real Estate Finance and Economics*, 17, 5–13. <https://doi.org/10.1023/A:1007783811760>
- Ping, J. L., Green, C. J., Zartman, R. E., & Bronson, K. F. (2004). Exploring spatial dependence of cotton yield using global and local autocorrelation statistics, *Field Crops Research*, 89(2–3), 219–236. <https://doi.org/10.1016/j.fcr.2004.02.009>
- Porta, M. (Ed.). (2014). *A Dictionary of Epidemiology*. Oxford University Press.
- Rechtin, M., Feldman, V., Klare, S., Riddle, N., & Sharma, R. (2020). Modeling and Simulation of COVID-19 Pandemic for Cincinnati Tri-State Area. arXiv preprint: 200606021. <https://doi.org/10.48550/arXiv.2006.06021>
- Ruktanonchai, N. W., Ruktanonchai, C. W., Floyd, J. R., & Tatem, A. J. (2018). Using Google Location History data to quantify fine-scale human mobility. *International Journal of Health Geographics*, 17, 28. <https://doi.org/10.1186/s12942-018-0150-z>
- Sardianos, C., Varlamis, I., & Bouras, G. (2018). Extracting User Habits from Google Maps History Logs. 2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM), Barcelona, 2018, 690–697, <https://doi.org/10.1109/ASONAM.2018.8508442>
- Schlosser, F., Maier, B. F., Jack, O., Hinrichs, D., Zachariae, A., & Brockmann, D. (2020). COVID-19 lockdown induces disease-mitigating structural changes in mobility networks. *Proceedings of the National Academy of Sciences of the United States of America*, 117(52), 32883–32890. <https://doi.org/10.1073/pnas.2012326117>
- Setti, L., Passarini, F., De Gennaro, G., Barbieri, P., Perrone, M. G., Borelli, M., Palmisani, J., Di Gilio, A., Piscitelli, P., & Miani, A. (2020). Airborne Transmission Route of COVID-19: Why 2 Meters/6 Feet of Inter-Personal Distance Could Not Be Enough. *International Journal of Environmental Research and Public Health*, 17(8), 2932. <https://doi.org/10.3390/ijerph17082932>
- Si, Y. L., Debba, P., Skidmore, A. K., Toxopeus, A. G., & Li, L. (2008). Spatial and temporal patterns of global H5N1 outbreaks. In ISPRS 2008: Proceedings of the XXI congress:

- Silk road for information from imagery: the International Society for Photogrammetry and Remote Sensing, 3-11 July, Beijing, China. Comm. II, WG II/1. Beijing: ISPRS, 2008. 69–74. International Society for Photogrammetry and Remote Sensing (ISPRS). http://www.isprs.org/proceedings/XXXVII/congress/2_pdf/1_WG-II-1/12.pdf
- Soures, N., Chambers, D., Carmichael, Z., Daram, A., Shah, D. P., Clark, K., Potter, L., & Kudithipudu, D. (2020). SIRNet: Understanding social distancing measures with hybrid neural network model for COVID-19 infectious spread. arXiv preprint: 200410376. <https://doi.org/10.48550/arXiv.2004.10376>
- Tsai, P. J., Lin, M. L., Chu, C. M., & Perng, C. H. (2009). Spatial autocorrelation analysis of health care hotspots in Taiwan in 2006. *BMC Public Health*, 9, 464. <https://doi.org/10.1186/1471-2458-9-464>
- Venkatramanan, S., Sadilek, A., Fadikar, A., Barrett, C. L., Biggerstaff, M., Chen, J., Dotiwalla, X., Eastham, P., Gipson, B., Higdon, D., Kucuktunc, O., Lieber, A., Lewis, B. L., Reynolds, Z., Vullikanti, A. K., Wang, L., & Marathe, M. (2021). Forecasting influenza activity using machine-learned mobility map. *Nature Communications*, 12, 726. <https://doi.org/10.1038/s41467-021-21018-5>
- Wang, Y., Wang, Y., Chen, Y., & Qin, Q. (2020). Unique epidemiological and clinical features of the emerging 2019 novel coronavirus pneumonia (COVID-19) implicate special control measures. *Journal of Medical Virology*, 92(6), 568–576. <https://doi.org/10.1002/jmv.25748>
- Wu, J. T., Leung, K., & Leung, G. M. (2020). Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: a modelling study. *The Lancet*, 395(10225), 689–697. [https://doi.org/10.1016/S0140-6736\(20\)30260-9](https://doi.org/10.1016/S0140-6736(20)30260-9)
- Yeshiwondim, A. K., Gopal, S., Hailemariam, A. T., Dengela, D. O., & Patel, H. P. (2009). Spatial analysis of malaria incidence at the village level in areas with unstable transmission in Ethiopia. *International Journal of Health Geographics*, 8, 5. <https://doi.org/10.1186/1476-072X-8-5>
- Yi, H., Ng, S. T., Farwin, A., Pei Ting Low, A., Chang, C. M., & Lim, J. (2021). Health equity considerations in COVID-19: geospatial network analysis of the COVID-19 outbreak in the migrant population in Singapore. *Journal of Travel Medicine*, 28(2), taaa159. <https://doi.org/10.1093/jtm/taaa159>
- Zarei, M., Rahimi, K., Hassanzadeh, K., Abdi, M., Hosseini, V., Fathi, A., & Kakaei, K. (2021). From the environment to the cells: An overview on pivotal factors which affect spreading and infection in COVID-19 pandemic. *Environmental Research*, 201, 111555. <https://doi.org/10.1016/j.envres.2021.111555>