

Merge'R'Us — Visualising Accessibility of Junior Colleges

Constance Tan
School of Information Systems, Singapore
Management University
hytan.2018@sis.smu.edu.sg

Fiona Wee
School of Information Systems, Singapore
Management University
fiona.wee.2018@sis.smu.edu.sg

Jasmine Toh
School of Information Systems, Singapore
Management University
jasmine.toh.2018@sis.smu.edu.sg

ABSTRACT

To investigate the effect of the mergers of some Junior Colleges on accessibility to students, the team created a Shiny application using R to aid in the visualisation of the distance and duration of areas from various schools. This paper discusses our implementation and the findings.

1. INTRODUCTION

The Ministry of Education of Singapore has called for the amalgamation and relocation of some schools, including primary schools, secondary schools as well as junior colleges. This is reported to be due to the fall in birth rate that led to lesser demand for education services and changes in geographical distribution of students. As a result, there is an adverse impact on accessibility for some students.

1.1 Project Objectives

In our project, we will build an application to investigate the effect of the new locations of the affected schools on the spatial accessibility to Junior College (JC) students.

Mainly, we wish to create an isochrone map and visualise the modelling results of Hansen accessibility model and Spatial Accessibility Measure. We will then analyse and interpret the output. We are also interested in identifying the affected residential areas which require more attention.

1.2 Project Motivation

Many families have been impacted in terms of accessibility due to the policy, with some moving houses or being inconvenienced. We wish to explore the extent of this as well as identify the areas impacted. Did the decisions meet its aim, or did it do more harm than good? We are curious about this, and regardless, if any area was impacted adversely, it would be good for measures to be implemented by the government to assist the families affected.

While the initial announcement was a few years back, this situation remains prevalent. There are still future plans for relocation, for example Yusof Ishak Secondary School from an area where demand for school places is falling to Punggol in 2021 to meet the growing demand [1]. While this is grounded, we wonder about the sustainability of moving schools everytime the demand falls. As the downward trend in Singapore's population growth continues, there might be more of such cases to come, which is a cause for concern.

At the same time, the rather recent lavish national expenditure on the building of the new Eunoia JC raised many eyebrows. Many questioned if the move was called for in light of the closures and mergers of the other schools [6]. This is an interesting issue we can look into on the side.

Moreover, this topic is close to home as we have members whose JC was merged. Hence, we wish to visualise and analyse the impact, and perhaps allow insight to be derived on the ideal locations of schools.

2. RELATED WORK

An existing application we reviewed was Balancing Urban Residential Plannings (BURP) [5]. This project looked into the impact of accessibility for families in residential housing after school mergers in Singapore. They calculated the distance of HDB flats to schools as part of their algorithms. Isochrone maps were also used to showcase the results. We also tried out their application and evaluated them as such

The application did well in displaying accessibility well displayed by travel time with public transport and Hansen Accessibility (based on duration or distance). They logically took public transport into account for accessibility and their use of kernel density is useful to see if HDB clusters are near a school.

The team also identified some rooms for improvements as follows. The application is unable to view the accessibility to schools given an address or location. Moreover, selecting another analysis method will uncheck the current checkbox. This is confusing for the user. A dropdown will be more intuitive as users will know that they are looking at only one specific model at a time.

We aim to address these concerns, and implement the future work mentioned for identifying ideal locations for schools.

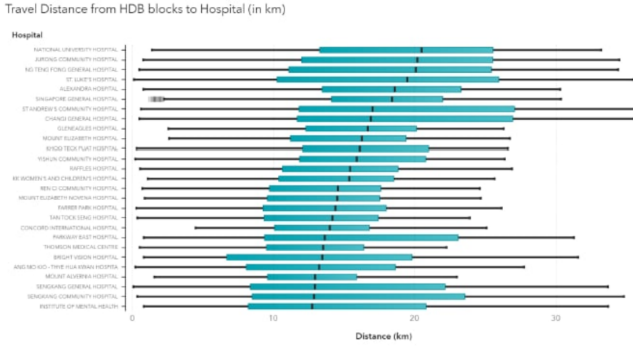


Figure 52: Boxplot of Travel Distance (in km) from HDB blocks to Hospital

Figure 1: Healthcare Study Boxplots

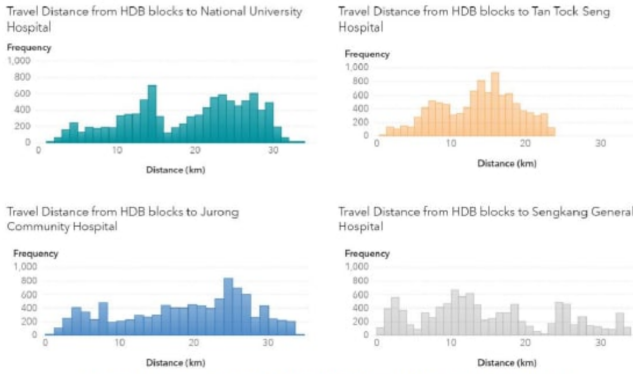


Figure 53: Comparison of Travel Distance between four hospitals

Figure 2: Healthcare Study Histogram

Another study, Is there a Space for SAS® Viya® in Healthcare? [4] had the aim of assessing if the locations of current (and upcoming) healthcare facilities are equally accessible by the senior population in Singapore. The exploratory data analysis they performed included box plots and histograms to compare the travel distance from HDB blocks to different hospitals as can be seen in the following figures. We find them useful to adopt in our application.

Methodologies applied include spatial point pattern analysis In the data exploration stage. They include First Order Distance-based and Second Order Distance-based functions and are used to understand the distribution of General Practitioner (GP) clinics and the interaction between the distribution of these clinics and that of the population. They also used gravity models to evaluate the accessibility of healthcare services to the senior population. Examples include Hansen Potential Model, Spatial Accessibility Measure (SAM) and 2-Step Floating Catchment Area.

3. DATA COLLECTION

The first dataset we used provides the JC information. We retrieved it using OneMap API and converted the response into “Jc.csv”. The data contains the fields for school name, postal code, latitude, longitude, X-coordinate, Y-coordinate, road name, address and region of each school.

The HDB property dataset “sg_zipcode_mapper.csv” is re-

trieved from Kaggle which contains a list of all postal code extracted from Singapore, including address, latitude and longitude and road name. The Kaggle dataset combines data from data.sg and the OneMap API to create a holistic dataset. We used HDB resale flat price from data.gov.sg to get the postal code and retrieve the proper address, latitude and longitude of each HDB from OneMap API. Python code is then used to extract the information [3].

We also procured the Master Plan 2014 Subzone Boundary (Web) data from data.gov.sg to obtain the region information for the JCs and the HDBs. Performing a join using the coordinates, we are able to merge it into the respective JC and HDB data.

The next data we have used is the General Transit Feed Specifications (GTFS) collected from singapore-gtfs repository [9]. It contains data of transit agencies file, distinct routes file which can be bus number or MRT line, trips file, frequencies file, stop_times file, geographic locations of the stops and calendar file which defines the service pattern.

The last dataset “Singapore.osm.pbf” is the Singapore’s OpenStreetMap PBF which contains the data of geographical area of interest. This PBF data would be used alongside the Open Planner (OTP) API (“otp.jar”) to help us in calculating the isochrones and the distance matrices. We will have to build the graph object by running the command `java -Xmx2G -jar otp.jar --build graphs/current` and proceed to run the OTP server by running the command `java -Xmx2G -jar otp.jar --router current --graphs graphs --server`. After the OTP server is up and running on our local server at `http://localhost:8080`, we will be able to proceed to our next step which is data generation.

To get the calculations, both the distance and duration, of HDB flats to desired JC, we generated distance matrices from each school to different HDB points using the OTP API by stating the router link and indicating the type of transport as walk and transit. We did this in data.RMD, which is used to do all our data generation. The time indicated for all OTP API generations are standardized at 7am, which is the peak hour where students travel to school. The distance matrices are then stored as a .rds file for future application of Hansen and SAM, which is covered in the next section.

4. METHODS

In our project, we implemented the methodologies covered by the case study, while also learning from the BURP application as reviewed in the previous literature review section. We also added other functionalities that were not covered.

Firstly, we implemented Isochrone mapping. This mapping is useful in measuring the accessibility of a JC based on how long it will take to travel from a certain point and time to the JC. The shorter the duration, the higher the accessibility of the school from the HDB point. Isochrone travel times were generated via the OTP API but instead of using the router link used for distance matrices, we made use of the isochrone link and indicated the same information [7].

Next, using Hansen’s Potential Model, we are able to mea-

$$P_i = \sum_j \frac{M_j}{d_{ij}^\alpha}$$

P_i = potential at School i

M_j = population for HDB point j , standardized as 100 for all

d_{ij}^α = distance between School i and HDB point j

Figure 3: Hansen Formula

$$A_i = \sum_{j=1} \frac{n_i}{p_j d_{ij}^\beta}$$

A_i = Accessibility of School i

n_i = supply for each School i

p_{ij} = population for HDB point j , standardized as 100 for all

d_{ij}^β = distance between School i and HDB point j

β = power of 2, following the theory of gravity in geography

Figure 4: SAM Formula

sure the accessibility of each JC, which is defined as the “opportunity which an individual at a given location possesses to take part in a particular activity or set of activities” with a certain cost (such as distance or time) [8]. In this case, the accessibility for each JC is calculated by looking at the potential that a person is willing to travel to the school based on the distance and duration from the HDB point to the school. The higher the value of Hansen, the higher the accessibility of the school from the HDB point.

Lastly, the Spatial Accessibility Measure (SAM) measures the accessibility according to the population of HDB points and supply of a JC [2]. The accessibility of a JC increased with higher population nearby and supply and decreased with longer distances/time from the HDB point to the JC. The higher the value of SAM, the higher the accessibility of the school from the HDB point.

For this project, we calculated both Hansen and SAM accessibility using the `ac()` function provided by the `SpatialAcc` package. This involves specifying the distance or duration matrix generated from Open Trip Planner API, and indicating power to 2 following the theory of gravity model in geography. As we are unsure of the population of each HDB point and the supply for each JC is difficult to gauge, we specified the population to be 100 and supply to be 10, and standardized this figure across all JCs. This means that we are evaluating the accessibility of each JCs purely based on distance or duration and more time needed or longer distance means less attractive.

The algorithmic implementations for the above methodologies are done in the `data.RMD` file. The result we got from generating the isochrones, and the calculations on Hansen and SAM using `ac()` function for each JC are stored as `.rds` files for use by the application. The specific implementation will be covered in the next section.

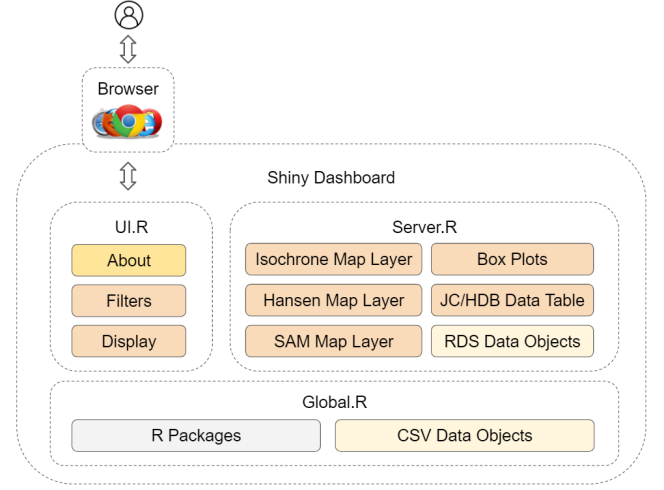


Figure 5: Application System Architecture and Overview

5. APPLICATION IMPLEMENTATION

The following figure shows an overview of how our application was built.

The main files driving the Shiny application dashboard that the user interacts with through their web browser are as follows.

5.1 Global File

Global.R lets us load the relevant libraries and datasets for the server and the UI, which allow us to perform various forms of analysis.

The R packages we used are as follows.

- Shiny features: shiny, shinydashboard, shinyWidgets
- Data manipulation: tidyverse which contains packages like dplyr and tidyr
- Spatial data handling for objects: sp
- Map: tmap
- Box plot: ggstatsplot
- Data table: DT

In our project, we used this file to create data objects to store our JC and HDB csv data. As not all the fields are relevant or in the format we want, we also perform data wrangling here. We filtered the data.frame to include columns that we want, and changed the format of the values to convert them from uppercase to camel case for more presentable display on the UI. We also set both data objects to the same coordinate reference systems, 3414, to coincide with that of Singapore's.

5.2 UI File

The UI.R file contains the dashboard layout and input components for our application, the functionalities provided of which are highlighted in the next section. Users can further interact with the models by applying the filters provided by UI.R, and the selection provided by the user will be communicated to the server, which will return the functionality

output to be displayed on the interface such as the different types of maps, the box plots and the data tables.

5.3 Server File

As mentioned above, the server supports the UI for rendering information. It observes input events by the UI, and when the user inputs their selections, Server.R will do the back-end work required, generating the information from the models to be displayed on the UI.

The core functionalities of our application are the isochrone and accessibility models to see the duration and/or distance. Upon selecting a JC or analysis type, the server will determine which is the analysis type chosen, and carry out the function to create the map layer based on the JC selected. The analysis for the JC is obtained by reading the respective RDS file generated from data.RMD as `sp SpatialPoints-DataFrame` object to be displayed on the map using leaflet. Using the pre-generated RDS files improve the performance of the application and allow the output to be shown in a shorter amount of time compared to letting the application generate the information on the spot.

At the same time, the legend is added by the server, as well as tooltips and markers if the user checks the specific checkbox to indicate that they want to see those additional information overlaid on the map layer created as well.

6. APPLICATION INTERFACE

Our application is a dashboard that allows users to view one of the following at a time: an interactive map, box plots for Hansen and SAM, as well as data tables displaying the information of JCs and HDBs respectively. The sidebar provides various options to be applied to customise and toggle their display of the component they wish to view.

To find out more about our application and how to use it, they may click the top right hand button with an information icon to reveal a notification menu. It also contains explanations on the various analyses used to help the user understand the application.

6.1 Interactive Map

On entering the application, users are brought to the “Interactive Map” tab first. A screenshot of this page is as follows.

In the map tab, users may make use of the following filters to see the results they want: region(s) for the JCs, JC, type of analysis, options to display JC and/or HDB points and finally they may customise the map theme.

The JC and type of analysis has to be chosen, and they are defaulted to the first options in the list. As the list might get quite long, they can make use of the region(s) input. By multi selecting the regions, it allows users to narrow their search for the JCs. This is convenient for users who wish to find JCs that are in their region. By default, all schools are displayed for selection.

For analysis type, there are 5 different selections to choose from:

- The ‘Duration (Isochrone)’ choice will display the Isochrone map of the selected JC. Areas within the red contour refers to the shortest time needed (within 15 minutes) to go to the selected JC thus the highest accessibility, while areas within the light sky blue contour refers to the longest time needed (within 90 minutes) to go to the selected JC thus lowest accessibility.
- The ‘Distance (Hansen)’ and ‘Distance (SAM)’ will display the mapping of Hansen Accessibility and SAM Accessibility based on the distance between HDB points to selected JC. HDB points with the darkest blue refers to the highest accessibility while HDB points with the lightest blue refers to the lowest accessibility.
- The ‘Duration (Hansen)’ and ‘Duration (SAM)’ will display the mapping of Hansen Accessibility and SAM Accessibility based on the duration needed to travel from HDB points to selected JC. HDB points with the darkest purple refers to the highest accessibility while HDB points with the lightest purple refers to the lowest accessibility.

Hovering over the HDB points for Hansen and SAM mappings will provide the specific accessibility value, address and distance or duration to the school.

After selecting a specific JC and analysis, the application will pull out the respective .rds file and display on the interface using leaflet.

There are several options for users to display other points by ticking the respective checkboxes according to their preference.

Selecting “Show all JC points” will display all JC points on the current map. JCs that are nearby will form clusters. The cluster size changes according to zoom level. Click on the cluster ball to view the schools within the cluster. Looking at the JC points can help to make the decision on amalgamation and relocation of schools as it allows easy comparison on the accessibility between schools.

Selecting “Show all HDB points” will display all HDB points on the current map. Hovering over the points will provide the user with the address of each point.

By checking the “Show chosen HDB point” checkbox, a search bar will appear for users to input a valid postal code. A tooltip of the address of the searched postal code will be displayed on the map upon valid search. This helps to identify the accessibility of a certain HDB pointed to the selected JC easily.

Users are also given the flexibility to choose the type of map they would like to work on by selecting the “Map Theme” dropdown list. Based on their selection, they may customise if they prefer light or dark themes, and if they want the map to contain other geographical features.

6.2 Other secondary features

Another tab that users may browse is the “Accessibility Measures” tab as follows.

On this page, they may see the box plot for Hansen and SAM side by side, based on the JC they select. From this, they can find out the distribution etc.

Next, users may take a look at the data tables containing information about the JCs and HDBs. They may perform sorting and search functions on these tables as shown in the screenshots below.

With reference to Figure __, they may filter the JCs shown by selecting the regions they want to see. Meanwhile, the HDB table may aid users to find the postal code of an address they search, such that they may input it in the “Show chosen HDB point” map selection input to show the location and details of that specific HDB block.

7. RESULTS

8. DISCUSSION

9. FUTURE WORK

Due to time constraints and other factors, our project still has room for improvement for better analyses. Our system could be further extended or refined as follows.

To improve our analysis, the following can be explored.

We identified that we can look into the demand and supply for each JCs and apply them to Hansen and SAM accessibility analysis to increase accuracy in the measurement of accessibility for each JC. We can do this by taking into account the school's reputation, number of intake, classrooms, teachers and space areas. Another aspect we can also look into the population for each HDB point by using the Singapore Residents by Planning Area/Subzone, Age Group and Sex from data.gov.sg. Although it does not contain any coordinates values, but it's Planning Area (PA) and Subzone (SZ) fields can be used as unique identifiers to georeference the Master Plan 2014 Subzone Boundary shapefile. We could also get the demand for HDBs from Urban Redevelopment Authority's Real Estate Information System portal (URA REALIS).

If it is possible, more recent data could be used, and/or all the different data used is to be standardised to the same period of collection. However, this may be hard to fulfil.

We could also include a spatio-temporal element to the analyses by comparing the maps between different years following any shifts in the locations of schools or otherwise. This can allow us to evaluate more clearly the effects of the merger decisions.

Currently, the results for the analyses types are given in the labels or tooltips found on the maps. These figures could also be included in the data tables, or we could have a tab for the summary statistics for different JCs or HDBs.

Other features that can be explored for better useability of our application as well include the following.

More customisation can be included. For example, taking the box plots tab, there could be a slider input to limit the values that the user wants to focus on, since there may be

outliers. There could also be toggle features so users can quickly choose which windows they want to expand or collapse, since there are quite a few and more might come. In particular, users can be allowed to choose if they want to view the distance or duration, or Hansen or SAM maps side-by-side, or to display all for their own comparison. They could even be allowed to drag the windows to create the layout they wish to view it in. For the map tab, radio buttons can be provided for the points to choose between show all JC or HDB points, or show only those in the selected region(s). In particular, the JC points display can be expanded more upon too as currently there is only one option. The application has also been coded in such a way that allows maintainability or easier update, as the one checkbox was coded as a grouped input, so more options may be easily added on in the future too.

It is also good to include other general information about the JCs or HDBs. In this case, the data we sourced from data.gov.sg includes the information for all types of schools, but did not seem to have data for all JCs so we did not implement it. This feature would be useful for users finding a JC or if there are families keen on finding and relocating to a new HDB such that it is close to a JC.

If the above is implemented, there would be many columns, and not all may be of interest to the user. We could include filtering options on the sidebar to allow them to select and deselect specific columns. It is also good to allow them to save these settings for future use.

We could also make it such that the row of the data table selected is the point the user wishes to view on the map. This might require some layout modification.

For more user friendliness, we could also add loading icons when the map layer is being loaded, or error messages when the postal code is invalid. We can also work on moving some of the messages in our information tab as tooltips for the specific input. This would allow for more user intuitiveness and avoid information overload or the case where users miss out on tips that would improve their experience as they have not read the notifications yet. Moreover, it is clearer to shift the position to somewhere nearer to the relevant components. For example, our application can suggest the user to visit the HDB Details tab to search for their postal code in the error message or nearby if they input an invalid code.

With respect to the performance of the application, there is some wait time or small lag when users toggle between analysis types or other features to allow the server to display the information. On top of adding the loading icon, the application can also be programmed to only clear the map layer when the next visualisation has been loaded.

It would also be good to provide information on the data sources used and the year each of them are obtained. We could also include modes of contact for them to reach out to us if they wish.

10. ACKNOWLEDGEMENT

We would like to thank Professor Kam Tin Seong for providing guidance and references to help us on our project.

References

- [1] Chia, L. 2018. MOE to relocate Yusof Ishak Secondary from Bukit Batok to Punggol.
- [2] Kalogirou, S. and Foley, R. 2006. Health, Place and Hanly, modelling accessibility to hospitals in Ireland.
- [3] Lee onemap api.
- [4] Lim, E. 2019. Is there a Space for SAS® Viya® in Healthcare?
- [5] Tan, P.C. et al. 2019. Balancing Urban Residential Planings (BURP), Visualising the Public Schools Accessibility to Residential Housings.
- [6] Tee, K. 2020. Look beyond state-of-the-art campuses for what matters in education.
- [7] Ten, D. 2018. Isochrone Maps with R and OpenTrip-Planner.
- [8] Walter G., H. 1959. Accessibility and Residential Growth.
- [9] Yang, Y. 2018. Singapore GTFS.