DenMap: A Dengue Surveillance System for Malaysia

Muhammad Rizwan¹, Sarat C Dass¹, Vijanth S Asirvadam², Balvinder S Gill³ and Lokman H Sulaiman⁴

¹Department of Fundamental and Applied Sciences, Universiti Teknologi PETRONAS, Seri Iskandar, 32610 Tronoh, Perak, Malaysia.
²Department of Electrical and Electronic Engineering, Universiti Teknologi PETRONAS, Seri Iskandar, 32610 Tronoh, Perak, Malaysia.
³Disease Control Division (Surveillance), Ministry of Health Malaysia, Complex E, Federal Government Administrative Centre, 62590 Putrajaya, Malaysia.
⁴Institute for Research, Development and Innovation, International Medical University Malaysia, Bukit Jalil, 57000 Bukit Jalil, Wilayah Persekutuan Kuala Lumpur, Malaysia.

muhammadrizwan312@gmail.com¹, sarat.dass@utp.edu.my¹,
vijanth_sagayan@utp.edu.my², drbsgill@moh.gov.my³,
LokmanHakim@imu.edu.my⁴.

Abstract. Dengue is the world's most rapidly spreading and geographically widespread arthropod-borne disease which needs proper surveillance and control strategies. Effective detection and containment of dengue has proven to be challenging due to the complex nature of interactions among the governing epidemiological factors as well as a number of environmental, climactic and societal attributes. An interactive approach of visualization and mapping of dengue cases on Google Maps is proposed here for the purpose of disease monitoring and surveillance. Our dengue visualization system is a web based application. It is developed using the Shiny package of R software along with HTML and JavaScript interfaces.

Malaysia has the two types of databases for dengue: the e-Dengue database for registered cases and the e-Notifikasi database for notified cases. The e-Dengue database has latitude and longitude information, obtained manually, for each registered dengue case. The visualization system uses those latitudes and longitudes to plot the registered cases on Google Maps. Notified cases do not have any latitude and longitude information but have a case location address for each case. Automatic batch geocoding of addresses is developed for notified cases. Subsequently, notified cases are rendered on Google Maps separately or together with registered cases to gain an understanding of the statistics of disease spread. Filtering options are developed according to either the notification or registration dates.

The dengue geocoding and visualization system, DenMap, is an interactive and easy-to-use application for producing maps of notified and registered cases rapidly. The maps display accurate locations of dengue cases and can be made to render other information of each case using map markers. This visualization and analysis system will not require any technical expertise for its usage.
1. Introduction

Dengue is the world’s most rapidly spreading and geographically widespread arthropod-borne disease. The World Health Organization (WHO) has reported a steep rise of annual dengue cases and numbers of countries reporting the disease since the 1950s. More recently, dengue transmission has been documented in Bhutan, Hawaii, the Galapagos Islands, Easter Island, Hong Kong and Macao, extensions of the usual geographical areas where the disease has been prevalent [1].

Furthermore, dengue epidemics were observed to be larger, more frequent and associated with more severe disease than they were in the past [2]. With a population of approximately two billion inhabitants, the risk of dengue is greatest in tropical and subtropical countries. With an estimated 120 million people travelling to these regions annually, travellers are at increased risk of contracting dengue [3]. WHO has estimated that over 50 million cases of dengue occur annually, of which up to 500,000 people develop severe dengue, resulting in hundreds of thousands of hospitalizations, and about 24,000 deaths [4]. Given the scope of the disease, the impact on health-care systems and economies of the countries involved is enormous. Transmission of dengue disease depends upon the interaction of mosquito, host and virus. Recent trends in Malaysia and worldwide reflect a rapid increase of dengue, mainly attributed to unprecedented population growth in urban centres of tropical countries, the increased and efficient movement of viruses in infected humans through modern transportation, the lack of effective mosquito control strategies, and climate factors. These factors have all contributed to the marked increase in epidemics of this disease [2]. Real-time detection of a dengue epidemic, therefore, becomes crucial for disease containment, prevention and control.

The main aim of surveillance in dengue is to monitor transmission and to facilitate prevention and control of the disease. In addition, assessment of disease severity, cost effectiveness of prevention programs and determination of burden of disease can be achieved [5]. Malaysia, and most other countries, uses a passive surveillance approach to monitor dengue cases. Passive disease surveillance systems depend on routine reporting of cases of diseases reaching health care facilities for treatment or service. This approach lacks sensitivity because of the low index of suspicion for diseases during inter-epidemic periods [6]. It is apparent that passive surveillance alone will not generate sufficient information needed for prediction of outbreaks [7].

Augmenting existing passive dengue surveillance data with a Geographical Information System (GIS) based visualization capacity and statistical modelling has recently provided a relatively effective early warning system for the disease. The forecasting of disease outbreaks using surveillance data has enjoyed a large degree of popularity in health related circles [8, 9]. GIS and statistical models have been extensively used to assist to predict dengue outbreaks, and have been increasingly incorporated into epidemiological research on infectious diseases [10, 11].

Under the Prevention and Control of Infectious Diseases Act 1988 (Act 342) in Malaysia, it is mandatory that all cases diagnosed as dengue fever (DF) or dengue haemorrhagic fever (DHF) and laboratory results for dengue are notified within 24 hours to the nearest health office. All clinical cases of dengue fulfilling the case definition for dengue and laboratory results are notified to the nearest district health office. Data on all dengue case notifications (notified cases) are entered into the e-Notifikasi database. Investigations of these notified dengue cases are carried out within 24 hours by trained environmental health officers, where all laboratory confirmed dengue cases (registered cases) and data on control activities are entered into the e-Dengue database. The e-Dengue database was established specifically for dengue case registrations and control activities. Dengue notification data was subsequently reported to state and national levels for surveillance and monitoring activities.

2. Motivation for System Development

For effective surveillance and control of dengue, it is important to understand the patterns and dynamics of disease spread over space and time [12, 13]. An effective surveillance program should also be feasible economically as affordability is one of the most important considerations when deploying adequate and sustainable monitoring systems. Notifications, already available in Malaysia and reported within 24
hours, serve as a real-time indicator of the severity of dengue occurrences, and thus, can be considered as a potential early warning indicator. Visualization of such data on maps enable health officials to obtain, analyse and understand real-time disease dynamics compared to the tabular or report forms. Displaying locations on maps serves as an exploration of data and can lead to cluster analysis of disease incidences. Generating maps for different time periods is helpful for understanding the disease progression over time.

A registered dengue case can be rendered on a map since its longitude and latitude information is available from the e-Dengue database. Notified cases, on the other hand, do not have latitude and longitude information in the e-Notifikasi database. However, each of these cases has a case location address. Latitude and longitude information, thus, can be obtained from this address by geocoding. Thousands of notifications are received every single month and geocoding all these cases manually one by one is not only tedious but almost impossible. Therefore, it is necessary to have an automatic batch geocoding software which could help to geocode large numbers of addresses automatically. Finally, rendering geocoded notified cases and registered cases on maps within designated time windows help healthcare officials to understand patterns of incidence and spread of dengue. Thus, the problems we faced are (i) the batch geocoding of case addresses, followed by (ii) displaying the geocoded address on spatial maps, and (iii) rendering both notification and registered cases, either together or separately, on maps together with other relevant information in map markers. We did not find a software in the literature that combined all three items (i-iii) as a streamlined sequence of tasks for epidemiological investigations.

3. Currently available tools
Many Geographic Information Systems (GIS) software and algorithms are available for the rendering and analysis of geographical data. For example, ArcGIS is a well-known GIS software used widely in epidemiology. It has a wide variety of features ranging from providing location maps for geographical data to spatio-temporal statistical analysis [14]. However, ArcGIS and many other GIS software are not open-source. In our development, we emphasize open-source platforms due to cost considerations and for use by all. Another problem with existing GIS software is that they require technical expertise to operate even for very simple outputs. Thus, a simple and specific application for surveillance purposes, with no technical expertise requirements, is more attractive to health officials compared to current GIS software. For our problem, we found that most existing GIS software do not perform batch geocoding of addresses. In our literature search, we found only Google Maps to provide facilities for batch geocoding through its Application Program Interfaces (APIs). These Google APIs, available in R and its packages (such as ggmap), allow for batch geocoding via the function geocode(). Since Google Maps APIs were used, geocoded locations were necessarily rendered on Google Maps. For this purpose, Google Maps has provided a free world base map without any restrictions on copyright. The integration of geocoding facilities in R and rendering geocoded data on Google Maps was achieved using other interfaces, namely, Shiny, JavaScript and HTML.

4. Implementation methodology using Shiny and R
Shiny along with embedded HTML and JavaScript codes were used to develop our application for dengue surveillance. Shiny is the web application framework of R statistical software. It is advantageous for multiple reasons: First, Shiny applications can be made to run locally, and therefore, there is no need to upload data on the internet, especially when there are privacy and security concerns, as in the case of addresses related to disease occurrences in this article. Second, Shiny enables the development of web applications that can integrate well with R scripts. This feature is useful in surveillance systems as R has a number of statistical analysis tools to detect clusters and trigger warning which can be incorporated easily to a Shiny web application. A previous system developed via Shiny was reported as the “DotMapper” [15], which only required to display existing longitudes and latitudes on maps. Thus, any freely available, open-source map provider, such as Leaflet, was adequate. However, in our case,
displaying on Google Maps is a necessity since geocoding was performed using Google Maps APIs via R.

5. Results and System Features

Our application is called DenMap which performs geoanalysis based on disease surveillance databases such as e-Notifikasi and e-Dengue databases. It consists of two main panels which appear as two tabs on top right corner; they are called the “Display on Maps” and “Geocoding” tabs. The “Display on Maps” tab is selected by default. The left part of the “Display on Maps” tab presents various user selections whereas the right part displays Google Maps or a table according to whichever is specified by the user on the left panel. The first field requires the user to select a file for display. It has all three options including notified cases only, registered cases only, and both notified and registered cases at the same time. When displaying notified cases only, the date option selects notified cases whose dates of notification fall within user specified date range.

![Figure 1](image.png)

**Figure 1.** Screenshot of the Visualization user interface of DenMap. Notifications (marked as green dots) on Google Maps indicate the longitude and latitude of potential dengue cases within the user specific date range specified on the left panel. Clicking on a green dot displays additional information on that notified case.

When displaying registered cases only, the date option selects registration dates that are within the user specified date range. Recall that registration dates are different from notification dates since there is a time lag for a notified dengue case to be registered as a confirmed case subject to laboratory results. In the case when both notification and registration cases need to be displayed, the date range selects all cases whose notification dates fall within the user specified date range. Notifications that turn into confirmed dengue cases at a later date (i.e., a registered case) is plotted in red whereas the rest (which did not turn into a confirmed case) are marked as green dots. The last option is valuable in understanding the mechanism of conversion of suspected (i.e., notified) cases into confirmed cases (i.e., registered) cases over space and time. For example, rendering such information in a map can help health
Figure 2. Screenshot of the Visualization user interface of DenMap. Notifications and registered cases (marked as green and red dots, respectively) on Google Maps indicate the longitude and latitude of potential and confirmed dengue cases, respectively, within the user specific date range specified on the left panel. Clicking on a green or red dot displays additional information on that case.

officials expend greater control efforts in localized regions that have the highest rates of conversions. Figure 1 shows the display of notifications from 01 January 2015 to 02 January 2015 on Google Maps based on geocoded data. Figure 2 shows the display of notifications and registered cases within the same date range. Summary statistics relevant to the display is also given on the left panel.

The second tab is the “Geocoding” tab. It is essential for acquiring latitude and longitude information for plotting based on case location or street addresses, such as for the e-Notifikasi database. The “Geocoding” tab provides an option to automatically geocode large batches of notified cases (Google maps API allows a maximum of 2,500 geocodes per day freely, and an option to geocode larger numbers with minimal charges). The first field under the “Geocoding” tab is the selection of the date range, which serves as a filter for any file which have a large number of cases with a large date range. The file loading button allows the user to select the file for geocoding. Once successfully loaded, pressing the Geocode button will start geocoding all those addresses that fall within the selected date range. Once geocoding is completed, the application updates the summary statistics which gives the percentage of addresses that were geocoded successfully. The user now has an option of whether to save the geocoded information permanently in the system or not. If so, the “Update DB” button will save the geocoded cases along with the latitude and longitude information in an .xls/.csv file within the application. Figure 3 shows the geocoding panel and how it appears when geocoding is being performed. For records which were not successfully geocoded during automatic geocoding process, there is an option to geocode them manually. The portion beneath the automatic geocoding panel is reserved for manual geocoding. The failed records can be geocoded one by one by editing their addresses manually. For this purpose, a list of records with unsuccessful geocoding attempts is displayed on the right side of the view when a user touches on “Refresh to get failed Notifications” button. This list will be displayed in form of a scrollable table. Users can select the notification number, address and postcode of a case from the table, place it in the respective fields in the right panel, and after proper editing, would able to get a satisfactory geocode for that case. After a successful manual attempt, the user is again given the
choice of saving these geocodes. Combining the tasks of automatic and manual geocoding, this tab serves as a key for effective and time saving geocoding tasks on large disease surveillance databases. Figure 4 shows summary statistic of the geocoding process earlier for the date 1st January 2015 as well as the table displaying all addresses that failed to be geocoded.

6. Deployment of the system
Principally this system has been developed for web usage for multiple users, but it can also be used as a single user on local machines. The benefit of using on personal local machines is the data safety, without removing any important information from data set. This application was developed mainly for health officials and surveillance staff in the Ministry of Health, Malaysia, but it can also be used by other researchers in a variety of geocoding and visualization applications.

7. Discussion and Conclusion
The proposed application generates maps for the registered dengue disease cases and notified cases. It is specifically designed for public health officials for monitoring and surveillance of disease in Malaysia. This application can also be useful to obtain an assessment of clusters and hot-spots of dengue cases by visualizing the cases on maps. It can show spatial as well as the temporal distribution of dengue cases easily, as well as give an idea how conversions from suspected to confirmed cases change over space and time. Most importantly, informatics based on notified cases which was not possible previously is
Figure 4. Table displaying the list of all addresses that failed to be geocoded for the day 1st January 2015. Summary statistic consisting of the number of addresses geocoded, the number of successfully geocoded addresses and the percentage of successfully geocoded addresses is given on the left panel.

now enabled via our system which provides automatic geocoding and rendering on maps. Since notified cases are required to be reported to the Ministry of Health within 24 hours, such quick and automatic geocoding and visualization capabilities allow for real time monitoring of dengue epidemics. The Shiny application is web based, and therefore, will also be able to be used on other devices enabling greater flexibility. The proposed system is easy to use and flexible. It has many features which are necessary to meet the requirements for the disease surveillance. All the development tools used here are open-source resulting in a very cost effective solution. It has all the flexibility to be upgraded and extended to a system for predicting future dengue outbreaks.

8. Acknowledgment
We would like to thank the Director General of Health Malaysia for his permission to publish this article. The authors would like to thank the Ministry of Higher Education, Malaysia, for financial assistance under the Exploratory Research Grant Scheme (ERGS/1/2013/STG06/UTP/02/02) and Fundamental Research Grant Scheme (FRGS/1/2016/STG06/UTP/02/5), as well as resource facilities provided by the Center for Intelligent Signal and Imaging Research (CISIR) at Universiti Teknologi PETRONAS.

9. Reference
[1] 2009 World Health Organisation. Dengue Guidelines for diagnosis, treatment, prevention and control. Geneva: World Health Organization
[2] Gubler D J 2002 Trends in Microbiology 10 100-2
[3] Halstead S B 2005 Emerging Infectious Diseases 11 740-1
[4] Gubler D J and Meltzer M 1999 Advances in Virus Research 53 35-70
[5] Teutsch S M and Churchill R E 2000 Oxford University Press, USA
[6] Gubler D J 2002 *Dengue Bulletin* **26** 96-106
[7] Gubler D J and Casta-Valez A 1991 *Bulletin of the Pan American Health Organisation* **25** 237-47
[8] Bowie C and Prothero D 1981 *International Journal of Epidemiology* **10** 87-92
[9] Allard R 1998 *Bulletin of the World Health Organisation* **76** 327-33
[10] Helfenstein U 1986 *Statistics in Medicine* **5** 37-47
[11] Catalano R and Serxner S 1987 *American Journal of Epidemiology* **126** 724-31
[12] Acharya B K, Cao C, Lakes T, Chen W and Naeem S 2016 *BMC Public Health* **16** 849
[13] Ling C Y, Greubner O, Cramer A and Lakes T 2014 *Geospat Health* **9** 131-140
[14] Haining R and Maheswaran R 2016 *Handbook of Spatial Epidemiology* ed Fagerberg J, Mowery D C and Nelson R R (Boca Raton: CRC Press) chap 4, pp 69-98
[15] Smith C M and Hayward A C 2016 *BMC Infectious Diseases* **16** 145