DEVELOPING SYSTEM ON URBAN HEALTH USING CROWDSOURCING METHOD: A CASE STUDY OF DENGUE OUTBREAK

Lyes Mokraoui and Norzailawati Mohd Noor

Department of Urban and Regional Planning, Kulliyah of Architecture and Environmental Design, International Islamic University Malaysia, Jalan Gombak, 53100, Kuala Lumpur, Malaysia
lyes.mokraoui@gmail.com and norzailawati@iium.edu.my

KEY WORDS: Aedes, GIS, civic engagement, volunteered geographic information, VGI

ABSTRACT
The alarming increase in the number of dengue fever (DF) cases in the world and particularly in Malaysia has made the disease one of the top priorities for its government. Over the years Malaysia has been collecting statistics about the disease and the vector itself. This study aims at collecting historical dengue occurrences and looking at the favorable conditions in which the mosquito thrives. The research aims at breaking the life cycle of the vector and the disease gradually by developing a system that would collect and process crowdsourced data and other factors i.e weather and moon phase to generate a dengue index. Moon phases, weather conditions, land use type and data from crowdsourcing are the study factors of this research. A collection of dengue cases numbers from the ministry's of health in Malaysia (MOH) has been compiled and correlated with some of these factors. Preliminary results have shown that dengue cases correlate with the moon phases, drought and raining seasons. This confirms that in Malaysia there is a relationship between the moon cycles and the mosquitoȏs night activity; this can be used as a short cycle factor alongside the raining-draught long cycles. We have further developed a method to monitor and break the life cycle of the vector of the dengue outbreaks in Malaysia through civic engagement followed by civic awareness by knowing high activity period the biting habit of the mosquito and by calculating the spatial correlation of non-temporal factors, intensive measures can be taken by the relevant authorities to forewarn the citizens. We believe that this system can be implemented in the immediate in the region saving lives and resources.

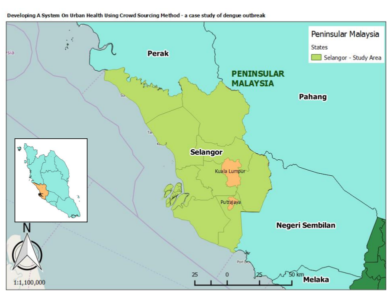
1.0 INTRODUCTION

Dengue Fever (DF) is the result of the effect of rapidly growing urban environment. It is a viral infection transmitted by Aedes mosquito female bite (WHO 2012). Incubation takes 3-14 days after the insect bite and has flu-like symptoms Dengue can be lethal if not dealt with on time (Bhatt et al, 2013). Dengue has four distinct serotypes and recovery from one provides immunity, but not against the three others (WHO 2015). With the alarming dengue fever outbreaks in the past years in the world (Murray et al, 2013) and Southeast Asia, Malaysia was not spared and suffers the burden of the disease. It is, in fact, the most alarming urban health related disease the country has been facing since its first outbreak in 1962. The urban health perspective which studies the negative effects urban environments have on the health of a population (Acevedo-Garcia, 2001) when promoting unhealthy lifestyles, such as ñconvenientò diets, sedentary behaviour, smoking, and the consumption of alcoholic substances (WHO, 2010); adding to the unhealthy lifestyle, making available man-made containers, transportation of recycled tires and other goods which accumulate water to form the perfect breeding ground for the mosquitoes are responsible for this rapid growth of the DF epidemic globally from 2 countries in the fifties (The Philippines and Thailand), to 9 by 1970 and reaching about 100 countries in 2014 (WHO 2015). DF does not discriminate; not only poverty and slums can contribute to DF as an urban health problem. Similarly, cities in the world with high level of wealth (Galea & Vlahov, 2005), Singapore (as a state city) and Putrajaya the new administrative capital of Malaysia with their modern urban planning are no exceptions to the outbreaks (Mulligan et al, 2012) even being free of slums but still accounting for high DF cases. In 1990, the Malaysian government has put in place a surveillance system at the hospitals and institute of medical research called ñDengue Virus Surveillance System" (DVSS) (Mudin, 2015). New methods of collecting data have emerged since the smartphone technological advances (Overeem et al, 2013). By bringing the internet to people wherever they, it has led to the emergence of popular data sharing methods named crowdsourcing. Studies have found that a social media system enabling could prevent DF outbreaks by developing a predictive surveillance system to forewarn the relevant authorities (Lwin et al, 2014). The mechanism of this system enables the data to be sent to servers using civic engagement. Using the crowdsourcing platform, it will enable citizens to relay data through an intuitive user interface. Existing research work was done in the state of Johor - Malaysia (Salawu et al, 2013) suggesting the use of a mashup web map to display and allowing manual Volunteered Geographic Information (VGI). In Malaysia, the number of mobile phone users could be used as a determinant factor in the vector surveillance through crowdsourcing. According to Malaysian Communications And Multimedia Commission (MCMC), as of 31st March 2014, there was a total number of 43,248,000 mobile phone subscribers throughout the country, with a penetration of 144.20 per 100 inhabitants (MCMC, 2015). The same study reveals that as per the survey date, about two-third (63.3%) of the mobile phone users access the Internet through their mobile phones (MCMC, 2015). These are encouraging statistics for our research framework. This research emphasis on the development of a system based on previous researches and historical dengue cases data as an attempt to predict where and when outbreaks would occur by combining five factors which affect DF outbreak incidence; rainfalls/droughts, temperatures, landuse, moon phases (Kampango et al, 2011) (Slooff, Verdrager, & Organization, 1972) and crowdsourced data (more of a surveillance tool) will be processed with Geographic Information Systems (GIS) techniques (Wen, Lin, Lin, King, & Su, 2006) (Saifur et al, 2013). Using these factors either spatial or temporal causing this disease and WHO steps to fight the vector (WHO 2015) we are developing a system to help reduce the outbreaks and the transmitting entity from spreading. We will subsequently establish a "Dengue Index" based on the known factors to predetermine short-term prediction of dengue outbreaks. The index will help local authorities to pinpoint the most probable dengue locations to use their resources in a systematic, timely and efficient manner. This index will use existing parameters influencing the vector's life cycle and biting habit (Wee et al, 2013). Prevention is possible by not giving access to egg-laying habitats to mosquitoes; removing potential man-made breeding grounds; applying the right insecticides at the right places; using protective households such as mosquito nets and "improving community participation". WHO has provided a list of preventive actions in order to control the transmission of the dengue virus (WHO 2015) by prohibiting the mosquitoes from accessing egg-laying environments, disposing of solid waste properly and timely, covering and disposing of domestic water storage containers, applying proper insecticides in a more effective manner, improving community participation in the DF prevention, using mosquito repellents, mosquito nets and active monitoring of the vector.

2.0 MATERIALS AND METHODS

2.1 Study Area

The study area is based on the data from Selangor state which is located along the west coast of peninsular Malaysia. The state boundaries are delimited by 2°35' and 3°60' N and between 100°45' and 102°00' E (Map 1). It is the most rapidly developing state in Malaysia (Abdullah & Nakagoshi, 2007). The choice of the state of Selangor which includes the federal territories of Kuala Lumpur and Putrajaya is due to the high number of dengue cases between the year 2000 and 2014 (Mudin, 2015) it has the highest rate of smartphone penetration is the country (MCMC, 2015). Selangor is the most populated in the country with a total of more than 6.1 million population excluding Kuala Lumpur and Putrajaya which is more than 20% of the overall population in Malaysia.



Map 1 - Location map of Selangor State in Peninsular Malaysia

2.2 Data Processing

Dengue data: Dengue data was retrieved from the Malaysian ministry of health website, it is available as pdf files which we compiled into a single Microsoft Excel file. Then using FME Software's Google Geocoder the data was converted into points objects stored as Esri' shapefiles. The limitation of these services is that free use is limited to 2500 addresses per day. The total of records were 11,430.

Moon Phase Data: The moon phase data was downloaded from "Time and Date" website. It consists of the dates in which each phase starts. By using Excel's function: WEEKNUM (date, first day of the week). The function would return the week of the month moon phase starting date. By this method, we can determine each week number for each moon phase.

Weather Data: Weather data used for this system are precipitation and temperature. It is suggested that the higher probability DF transmission occurs at temperatures equal to 28 °C (Chen & Hsieh, 2012). Since the temperatures in

Malaysia are in average above the threshold, we would assume that this variable will be set aside as the country has a favorable breeding and mosquitoes activity temperature. The effect of the rain in creating the breeding sites and the survival of the mosquito larvae is proven as described in the literature review. It takes about 25 days for the young mosquito to be actively feeding.

Land Use Data: Landuse data for Selangor was requested from the Malaysian Department of Town Planning (Jabatan Perancang Dan Desa) as they are the custodians of the land use data in the country. All the dengue data provided by the ministry of health are located in housing areas.

Crowdsourced data: The human data factor that would be relayed through VGI is the fifth variable used to generate the index. There is a need to facilitate the citizens-authorities communication flow in terms of epidemiology information. The gap in relaying information to the right agencies. The developed crowdsourcing App will be used as an extension for the citizens volunteering to be involved to reduce DF burden. It is important to note that the App will not be successful unless it is built with "simplicity" in mind. The collected data itself is very simple, the main factors to be collected are: Location (coordinates), biting activity, uncollected rubbish, stagnant water with larvae and the knowledge if there is/are dengue cases in the area.

3. SYSTEM DEVELOPMENT:

The developed system will have as objectives, gathering the data in an automated manner; processing the data, computing the Dengue Index and displaying the data on a web map (a heatmap). The data preparation will be organized in two stages:

3.1. Pre-System Development: This stage, includes gathering data from different sources and filtering them. Once restructured it would be uploaded into the main repository database at a final step to this stage pre-System Development. This is the most tedious and time-consuming stage.

3.2. During the System Development: At this stage, the initial data would have been uploaded into the main repository. Data feeding into the database will be done on a scheduled basis using FME Desktop/ Server from Safe Software. Since the data architecture has already been set in the previous stage, this would only be a matter of making updates of the new coming data from weather, moon phase and Crowdsourcing. All data will be updated on a daily basis.

3.3 Components of the System: The system after development will have four main components (Figure 1):

i. Servers: a Storing the data: The data will be stored in Postgres/PostGIS databases. The database software is open source and it has a high number of users worldwide. b. Publishing the maps on the web: The same server will also host web technologies which are PHP to interact with the database; Geoserver and OpenLayers to display maps on the web. PHP will also serve as a relay mechanism to update the Apps data into the database.

ii. Desktop Software: QGIS desktop software will be used to display and author the maps layers before being uploaded into PostGIS database and Geoserver. FME Desktop will be the main engine for scheduling the daily or weekly data updates.

iii. The Apps for VGI: The Apps data represent only 20% of the entire dengue index computation.

iv. Web Component: The heatmap will be displayed on a website accessible by everyone.

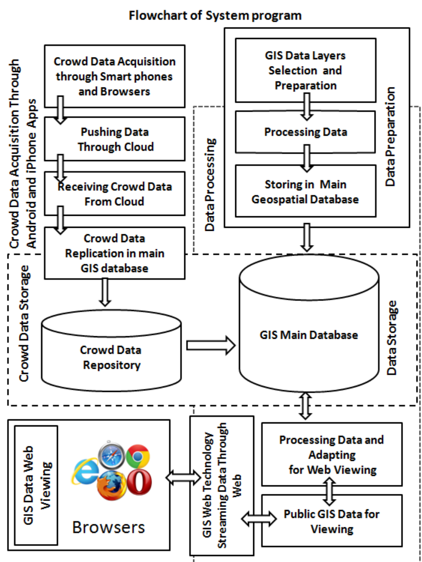


Figure -1 - Flowchart of the System

4. DENGUE INDEX

The outcome of this research would be to generate the scalable dengue probability outbreak. The Index will be a key measure to quantify the severity of the dengue outbreak prediction. The proposed Dengue Index would be calculated based on the five factors. The 5 factors have been classified into: Temporal, Spatial and temporal-Spatial. The index formula would be based on the calculation from all active temporal and spatial factors which would make a scale from 0 to 100. Where zero would be the indicative of a no risk of contracting DF or any A. aegypti based disease and 100 being the most critical in the event if all the factors are available.
The proposed formula would give a 20% credit to each factor. There are fixed parameters and temporal/spatial parameters (Figure 2).
The index will be generated based on the following formula:

□= □□+ □□+ □□+ □ □+ □□

Where:
Lu: Spatial location with a value of 20 on residential and 0 for other locations
Pr: Spatial location during rainfall where raining period exceeds 25 days with a value of 0 during periods of rain and 20 during periods of drought.
Te: Temperature component with a value of 0 when the temperature is below 28° C and 20 when temperatures are 28° C and above.
Mp: Moon Phase value will take the following values:
New Moon Phase:0
First Quarter Phase:10
Full Moon and Third Quarter:20
Cs: Crowdsourced Data, values will vary as follow:
No data: 0
Uncollected rubbish: 5
Mosquito Bites: 5
Mosquito Larvae: 5
Known Dengue case in the area: 5
Once the index calculated and localized spatially, it will be easy to correlate its spatial repartition. A heatmap layer using the point density method will be used to visualize the Index density of point features using Web mapping technology. Using a density map compared to buffered areas around the points of dengue cases location is that the vector is very complex and can fly great distances in a day. By creating the index maps in such way, we will be following a common path or the alley tracing the vector's direction.

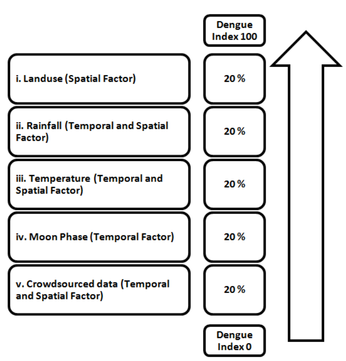


Figure 2- Dengue Index - Generated by the System

5. FINDINGS AND ANALYSIS

The analysis of the dengue cases data to establish the correlation with the moon phases was done based on two variables; the mean of the number of cases and the standard deviation. The statistics were generated using FME Desktop software grouping the data by weeks of the year then exporting the results into an Excel format. The histograms were then generated using Excel software. The findings show that in both graphs representing the temporal repartition of the DF cases data with the moon phases either for the mean values (Figure 3) or the standard deviation (Figure 4) - in this example data representing DF cases for years 2014 - a correlation in which the new moon phase in most of the time is the phase with the lowest number of DF cases or the period of time with a break before they start increasing abruptly. This suggests that most of the infected bites take place either during the full moon or during the third quarter of the moon phase.

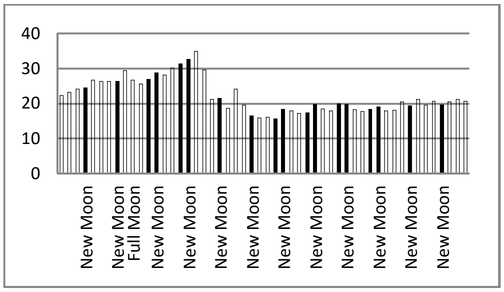


Figure 3 Dengue Cases Correlation Mean - Moon Phase Selangor - 2014 Data

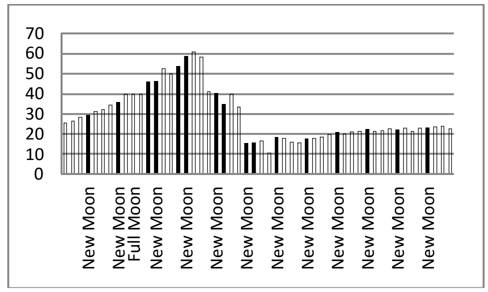


Figure 4 - Dengue Cases Correlation Standard Deviation - Moon Phase Selangor - 2014 Data

6. RECOMMENDATION AND FUTURE OUTLOOKS

The index takes in consideration spatial and temporal parameters involving DF outbreaks. Such index will not be dependent on the number of dengue cases but instead, it will rely on the parameters creating the favorable breeding habitats for the mosquito larvae. It will also highlight the time of the month when the probabilities of contracting dengue are the highest. By displaying the index as a heatmap on a web map, the local authorities will be able to pinpoint the areas with a higher dengue index, hence notifying the authorities about the outbreak before happens. In addition, to developing the system and the algorithm to compute the index and displaying it, this study suggests that the biting habit of the vector increases during full moon and the third quarter moon phase. We would, however, advise the Malaysian ministry of health to adjust the way data is compiled and follow the moon calendar weeks rather than the Gregorian calendar weeks. In this way, we could confirm the correlation between the DF cases and the moon phases.

7. CONCLUSION

There is a need to facilitate the citizens-authorities communication flow in terms of epidemiology information. The gap in relaying information even voluntarily makes the dengue vector a winner. The proposed crowdsourcing Apps to be developed is an extension to the citizens to be involved in the fight. The dengue index would be the base for measuring the probable DF outbreak. In addition to well-known dengue factor such as temperature and rainfall, the system itself would take in consideration two new factors; the moon phases and the human data factor that would be relayed through VGI. The importance of this research is to apply a method not only in Malaysia but this could be applied directly in other tropical and subtropical countries. The proposed index will be scalable depending on the region where it is used. It is also important to note that such research project will not be successful without the help of the governments in using the media to encourage the people to download and use the App for their own benefit. The developed system could be used to forewarn the outbreaks three weeks before their incidence as an attempt to reduce the dengue occurrences every year in Malaysia. The DF reduction rates would only be verified if such research project is opened to the public at a large scale. We need to take into consideration the fact that the users who would like to help will come from various backgrounds. Making a complicated App would discourage them from helping. Once the system is proven to be beneficial, it could be easily replicated in the rest of the states in Malaysia and even in other countries suffering from the heavy dengue burden.

8. ACKNOWLEDGEMENT

The authors greatly acknowledge the International Islamic University of Malaysia and Ministry of Health, Malaysia for providing invaluable respective data used in this study. Authors sincerely thank all referees for their suggestions to improve the manuscript.

9. REFERENCES

Abdullah, S. A., & Nakagoshi, N. (2007). Forest fragmentation and its correlation to human land use change in the state of Selangor, peninsular Malaysia. Forest Ecology and Management, 241(1-3), 39ï 48.

Acevedo-garcia, D. (2001). Zip Code ï Level Risk Factors for Tuberculosis ¿ Neighborhood Environment and Residential Segregation in New Jersey , 91(5), 1985ï 1992.

Bhatt, S., Gething, P. W., Brady, O. J., Messina, J. P., Farlow, A. W., Moyes, C. L., é Hay, S. I. (2013). The global distribution and burden of dengue. Nature, 496(7446), 504ï 7.

Chen, S. C., & Hsieh, M. H. (2012). Modeling the transmission dynamics of dengue fever: Implications of temperature effects. Science of the Total Environment, 431, 385ï 391.

Galea, S., & Vlahov, D. (2005). URBAN HEALTH ¿ Evidence , Challenges , and Directions.

Kampango, A., Cuamba, N., & Charlwood, J. D. (2011). Does moonlight influence the biting behaviour of Anopheles funestus? Medical and Veterinary Entomology, 25(3), 240ï 246.

Lwin, M. O., Vijaykumar, S., Fernando, O. N. N., Cheong, S. A., Rathnayake, V. S., Lim, G., é Foo, S. (2014). A 21st century approach to tackling dengue: Crowdsourced surveillance, predictive mapping and tailored communication. Acta Tropica, 130, 100ï 107.

MCMC. (2015). Hand Phone Users Survey 2014. Communications, 48.

Mudin, R. N. (2015). Dengue Incidence and the Prevention and Control Program in Malaysia.

Mulligan, K., Elliott, S. J., & Schuster-Wallace, C. (2012). The place of health and the health of place: Dengue fever and urban governance in Putrajaya, Malaysia. Health & Place, 18(3), 613ï 620.

Murray, N. E. A., Quam, M. B., & Wilder-Smith, A. (2013). Epidemiology of dengue: Past, present and future prospects. Clinical Epidemiology, 5(1), 299ï 309.

Overeem, A., R. Robinson, J. C., Leijnse, H., Steeneveld, G. J., P. Horn, B. K., & Uijlenhoet, R. (2013). Crowdsourcing urban air temperatures from smartphone battery temperatures. Geophysical Research Letters, 40(15), 4081ï 4085.

Saifur, R. G. M., Hassan, A. A., Dieng, H., Rawi, M., Salmah, C., Saad, A. R., & Satho, T. (2013). Temporal And Spatial Distribution Of Dengue Vector Mosquitoes And Their Habitat Patterns In Penang Island, Malaysia. Journal of the American Mosquito Control Association, 29(1).

Salawu, A., Safie, M., Bello, A. A., & Bernad, S. C. (2013). Geo Visualization of Disease Outbreak Using Geospatial Crowd Sourcing Web Application. International Journal of Research and Reviews in Applied Sciences, 16(3), 370ï 376.

Slooff, R., Verdrager, J., & Organization, W. H. (1972). Anopheles balabacensis balabacensis baisas 1936 and malaria transmission in South-Eastern areas of Asia. Geneva ¿ World Health Organization.

Wee, L. K., Weng, S. N., Raduan, N., Wah, S. K., Ming, W. H., Shi, C. H., é Lim, L. H. (2013). Relationship between rainfall and Aedes larval population at two insular sites in Pulau Ketam, Selangor, Malaysia. Southeast Asian Journal of Tropical Medicine and Public Health, 44(2), 157.

Wen, T. H., Lin, N. H., Lin, C. H., King, C. C., & Su, M. D. (2006). Spatial mapping of temporal risk characteristics to improve environmental health risk identification: A case study of a dengue epidemic in Taiwan. Science of the Total Environment, 367(2-3), 631ï 640.

WHO. (2010). Why urban health matters.

WHO. (2012). WHO What is dengue and how is it treated. World Health Organization.

WHO. (2014). WHO | Vector-borne diseases. World Health Organization.

WHO. (2015). WHO | Dengue and severe dengue. World Health Organization.

Weather data for Malaysia extracted from the National Oceanic and Atmospheric Administration's website. http://www.noaa.gov/weather

Moon Phases data for Kuala Lumpur - Malaysia extracted from Time and Date website http://www.timeanddate.com/

Dengue Cases data for Selangor state extracted from Ministry of Health Malaysia website , http://www.moh.gov.my/