Clustering of Domestic Locations in Layers for the Purpose of Breakout Prevention of Diseases

Varin Chouvatut[1], Ekkarat Boonchieng[1], Waraporn Boonchieng[2]

[1] Department of Computer Science, Faculty of Science, Chiang Mai University, Thailand
[2] Faculty of Public Health, Chiang Mai University, Thailand
varinchouv@gmail.com, ekkarat.boonchieng@cmu.ac.th, waraporn.b@cmu.ac.th

Abstract

The aim of this research is to help support Provincial Public Health Offices (PPHOs) by clustering geographical coordinates in terms of latitudes and longitudes and comparing each cluster's density in order to prevent mosquito-born diseases. In this research we propose using the hierarchically spatial clustering approach with two layers of clustering. We cannot use pre-existing clustering algorithms because they do not fit our purpose due to the naturally breeding area of common house mosquitoes. Thus, a modified method with an additional strict constraint of k-means and k-nearest neighbors were applied to meet this goal. Additional constraint to k-means clustering in the first layer will cause some of the coordinates to un-cluster. Finally, all coordinates must be clustered again to a second layer. However, the constraint of mosquito's breeding area must be retained thus some falsely alarmed coordinates must be filtered out.

Keywords: Disease breakout, Geographical coordinate, Hierarchical clustering, Mosquito breeding

1 Introduction

Recently, the Provincial Public Health Offices (PPHOs) who's Ministry of Public Health in Thailand has sent officers out to locations where they use fogging machine to blow chemical substances around human households and drainage grooves. They target areas where mosquitoes most likely spawn during the rainy season in order to prevent the mosquitos from breeding. Without this protection some areas may become vulnerable to disease outbreaks. These diseases include Chikungunya and Dengue Hemorrhagic fever. To process on this outbreak protection in a large area, PPHO needs to spend a large fleet of officers who incur high travel expense. But the PPHO does not need to waste its money for this process if it knows or can limit the vulnerable areas instead of randomly selecting some areas.

Data mining is a strategy for extracting or mining knowledge out of a raw big data [10-11]. Data mining has a simple and neat method for clustering large data called k-means whose original algorithm can be found in [1-2]. Researchers in [1] applied k-means algorithm to group students into different clusters and predict weather or not s student will take the computer proficiency test and fail. The prediction results can be used to warn a professor to take better care of some groups of students. Researchers in [3] analyzed that many clustering algorithms categorize each instance into a group and some algorithms like k-means categorize all instances in data set into one of the clusters. This means that, the k-means algorithm will not allow any ungrouped exceptions. According to [4], researchers demonstrated that the popular k-means algorithm, an unsupervised method for data analysis, can be modified and thus used for automatically detecting road lanes from the Global Position System (GPS) data. Researchers in [5] proposed an approach for classifying and clustering the ABAC policies. This method allowed some scale of flexibility where the k parameter was controlled. Researchers in [6] used a modified version of k-nearest neighbors for constructing a spatial network V*-diagram. They defined spatial regions to be used with k-nearest neighbors as the Voronoi diagram. The regions were the query points which might move with no need to change the query answer.

In this research, we propose methods for clustering areas that may be considered a vulnerable area where common house mosquitoes may breed. Since clustering result using solely the k-means method contains some overlapping regions, we need to modify it further to improve area grouping and clustering. Then k-nearest neighbor algorithm with local distance-based filtering strategy is applied from [7]. This is used for further clustering because its original version was used for clustering in [5]. The input information is geographic coordinate system composing of latitude and longitude of locations where mosquito larvae have been found. Once the regions were clustered, the density of each cluster will then be measured and sorted for the purpose of priority consideration. Thus, the areas with

*Corresponding Author: Waraporn Boonchieng; E-mail: waraporn.b@cmu.ac.th
DOI: 10.3966/160792642020112106011

higher priority will be processed for common house mosquito's breeding prevention first. Consequently, PPHOs would then work on the process of mosquito's breeding prevention more quickly and efficiently.

2 Data and Methodology

The techniques we used to input and process data are described as follows.

2.1 Input Data

We used geographical locations obtained from The Center of Excellence in Community Health Informatics (COEiCHI), Chiang Mai University, Thailand funded by Chiang Mai University, Thailand.

The geographical locations are the areas where mosquito's larvae have been detected. These data were reported online in real-time through an official application developed by COEiCHI [12-13] and sponsored by Thai Health Promotion Foundation of Chiang Mai Provincial Public Health Office, a government organization of Thailand. The locations that were chosen are households located in five districts of Chiang Mai, a province in the north of Thailand. The locations of each household in 5 districts used in the experiments are featured in Table 1. Table 2 is an example of input .xls file showing details of each household location in Saraphi district including address, (some of the digits are replaced with 'x' for confidential privacy), Moo, sub-district, district, living time duration (in years), and geographical coordinates (some of the digits are replaced with 'x' for confidential privacy). Note that Table 2 has been separated into two parts, Part 1 and Part 2, where Table 2 (Part 1) includes sub-district and district of each address while Table 2 (Part 2) includes living time and geographical coordinates of each address. In this experiment, there were 744 households in the Saraphi district in total. We chose the Saraphi district because of the large number of households located there. In addition, these data are related to dengue fever detection using Long Short-term Memory Neural Network [14].

2.2 Hierarchically Spatial Clustering Method

In this paper, we choose hierarchical or layered clustering method to categorize or group all available locations. Geographical coordinates in terms of latitude and longitude vary in a large area such as a district. The geographical locations used were in this paper are real world locations. Some of these locations have large distances (many kilometers) between them, even ones located in the same district.

The main advantage of applying hierarchical clustering where the first layer is an applied version of k-means clustering and the second layer is k-nearest neighbor classification is that we cluster a certain number of locations based on distances from centroids of k clusters as initial then any location considered too far from any centroid of k clusters will be allowed to be categorized later in the second layer with higher flexible constraint. Consequently, we do not need to set high number of clusters, k, initially. With large number of k, convergent time for calculating a steady mean of each group out of k groups can even increase if each group has large number of members or locations.

Table 1. Real locations from five districts in Chiang Mai, a province in the north of Thailand

Data Set	District	The Number of Households
1	Saraphi	744
2	Hang Dong	118
3	San Pa Tong	205
4	Doi Saket	24
5	Doi Tao	54

Table 2 (Part 1). A part of an example Microsoft Excel file in .xls format keeping details of households in Saraphi district with a report of mosquito's larvae found

Address Number	Moo	Sub-district	District
1x	5	Tha Kwang	Saraphi
2xx	1	Nong Phueng	Saraphi
2xx/1	1	Nong Phueng	Saraphi
4x/2	11	San Sai Mahawong	Saraphi
4x/1	11	San Sai Mahawong	Saraphi

Table 2 (Part 2). A part of an example Microsoft Excel file in .xls format keeping details of households in Saraphi district with a report of mosquito's larvae found

Address Number	Moo	Living Time (years)	Geographical Coordinates	
			Latitude (°)	Longitude (°)
1x	5	26	18.9024xx	100.523xx
2xx	1	35	18.474230xx	99.013432xx
2xx/1	1	27	18.474341xx	99.01340xx
4x/2	11	59	18.618032xx	98.959785xx
4x/1	11	52	18.618664xx	98.959321xx

2.3 The First Layer

An automatic clustering method is k-means, given the number of clusters, k, locations from the input data can be categorized into k groups based on distances between the considering datum and the current centroid of each cluster [8]. The considering datum will be categorized to the group or cluster whose centroid is closest to the datum. From the given geographical locations as input data, we employed the clustering method with an applied version of the original k-means algorithm as follows.

2.3.1 Randomization of k Locations

Randomize k locations from available geographical locations provided as input from Section 2.1. These k locations will be set as centers or centroids of k clusters. Note that, at this step, we did not randomize any pair of latitude and longitude value, which is not a real location available (typically done by the original method), instead we randomize locations from the given input locations only. The purpose of this choice is mainly to limit scope of location disallowed to be initially set too far from the real coordinates of interest. In addition, to provide further adjustment a good initial set of centroids.

2.3.2 The Rest Locations

Categorize the rest locations which are not chosen as any cluster's centroid in the first step (Section 2.3.1) to a cluster. To categorize a location to a cluster, we calculated distances between the location which is defined as latitude and longitude values and each location of centroids of k clusters. Then, the location will be categorized to the closest cluster amongst k clusters whose distance between the location and the centroid of the defined cluster is shortest comparing to the other k-1 clusters. That is, for each location, the cluster closest to the considering location, $d\_k_{shortest}$ must be calculated as

$$d\_k_{shortest} = \min \{d\_k_i \mid i = 1..k\} \tag{1}$$

where d_k_i is the distance in kilometers measured from centroid of the ith cluster from k clusters to the considering location and can be computed in terms of latitude and longitude values using (4) and (5).

After $d\_k_{shortest}$ was obtained, the considering location will be classified or categorized to the ith cluster whose d_k_i is the shortest distance amongst k clusters. However, we added a constraint to this step of categorization based on measured distances in the way that any location whose shortest distance, $d\_k_{shortest}$, is longer than a predefined threshold will not be counted into the ith cluster. This threshold will be corresponding to the radius defined by PPHOs of naturally breeding area of the common house mosquitoes. Here, we set the threshold radius to two kilometers for experiments as default. Consequently, after convergence of the algorithm reached, there will generally be some locations left uncategorized to any cluster due to this additional constraint. Then, such locations will be further pushed to the second layer of hierarchical clustering to categorize them completely.

2.3.3 Centroid Update

Average or mean value of each of k clusters will be computed and defined as an updated centroid of the cluster. Thus, after this step, the new location in terms of latitude and longitude values obtained from average calculation as (2) will be the new centroid of that cluster.

$$Clat_i = \frac{\sum_{j=1}^{n} lat_j}{n}, \quad Clong_i = \frac{\sum_{j=1}^{n} long_j}{n}; \tag{2}$$

where $Clat_i$ and $Clong_i$ are latitude and longitude, respectively, of the new centroid of ith cluster, lat_j and $long_j$ are latitude and longitude values of jth location in ith cluster when j = 1..n representing the number of locations, n, in ith cluster.

2.3.4 Repetition

Repeat Sections 2.3.2 and 2.3.3 until ($Clat_i$, $Clong_i$) for i = 1..k are either unchanged or insignificantly changed (in other words too small amount of change has been detected).

2.4 The Second Layer

After applying k-means algorithm as the first layer for clustering data, there will still be some geographical locations cannot definitely be clustered in one group since they are out of range of radius of any cluster defined by k-means.

Since the areas obtained from Section 2.3 are circular shapes and for consistency with the purpose of avoiding bias in clustering, we should use algorithm providing the similar shape of the considered areas, thus further clustering the rest incompletely clustered areas in the second layer of clustering should be done with the clustering method based on circular areas. We thus chose k-nearest neighbors. The k-nearest neighbor method is an algorithm for classifying or categorizing data based on the number of nearest members (breeding locations) from each cluster's centroid [9]. However, an additional constraint for this layer is still be distances between the contiguous locations, that is, any two nearest locations must not be further than a threshold distance, e.g. two kilometers. Finally, after this step, all geographical locations will belong to a cluster.

2.5 Filtering Unexpectedly High Density of a Small Cluster

Although the second layer of hierarchical spatial clustering can group the other locations that were abandoned from the first layer, there can be some clusters with only single location (which should even not be called a cluster) and some clusters with a considerably small number of locations. These kinds of clusters should no longer be considered a highly risky area and thus should be filtered out. Thus, any cluster with small number of locations but with high density value calculated will be removed from being considered as a highly risky area of mosquito's breeding.

2.6 How to Measure Distances

In order to use distance as a base information for making decision if a specific location should belong to which cluster, the distance between the location and the recent centroids (a recent centroid is for example the recent center with respect to the newly calculated mean of a cluster) of all clusters. Since the latitude and longitude of each location are in the unit of degrees (°), lat_d and long_d, the location must be converted into radians, lat_r and long_r :

$$\text{lat\_r} = \text{lat\_d} \times \frac{\pi}{180},$$
$$\text{long\_r} = \text{long\_d} \times \frac{\pi}{180}. \tag{3}$$

Then, the distance d_m(g_1, g_2) in the unit of miles between a pair of geographical locations g_1 = (lat_1, $long_1$) and g_2 = (lat_2, $long_2$) where lat_1 and $long_1$ are the latitude and longitude values, respectively, of location g_1 and lat_2 and $long_2$ are the latitude and longitude values, respectively, of location g_2 can be calculated from

$$\text{d\_m}(g_1, g_2) = \text{acos}(\sin(lat_1) \times \sin(lat_2) + \cos(lat_1) \times \cos(lat_2) \times \cos(long_1 - long_2)). \tag{4}$$

To convert d_m(g_1, g_2) into kilometers, d_k(g_1, g_2), we used

$$\text{d\_k}(g_1, g_2) = \text{d\_m}(g_1, g_2) \times 60 \times 1.1515 \times 1.609344. \tag{5}$$

2.7 Density of Common House Mosquito's Breeding Areas

Since the areas calculated from the clustering method we used in Sections 2.3 and 2.4 are in circular shapes, the population density, D, of N common house mosquito's breeding areas in a cluster whose radius measured from the cluster's centroid equals R kilometers obtained from Section 2.4 can be computed from

$$D = \frac{N}{\pi \times R^2}. \tag{6}$$

Once the densities of breeding areas of all clusters are calculated, they will be sorted in descending order to find out which cluster (or region) are in the highest risk of high rate of mosquito's breeding and probability of facing the diseases caused by common house mosquitoes as carriers. Thus, these regions must be taken in action at the first place.

3 Results and Discussion

To display geographical coordinates clearly and sensibly, Google Map API is used for all experimental demonstrations in this section.

3.1 Experimental Results from the First Layer of Hierarchical Clustering

An example result obtained from the first layer is as Figure 1 below. After applying the first layer with k = 10 clusters and threshold radius = 2 kilometers, there are still some locations displayed in red which cannot be clustered due to exceeded distance from any cluster's centroid. From Figure 1, according to the constraint of distance spread around out of each cluster's centroid, ten clusters existed from the first layer with densities shown in Table 3 where density column is sorted in ascending order. Note that Table 3 has been separated into two parts, Part 1 and Part 2, where Table 3 (Part 1) includes the number of breeding locations and radius of each cluster while Table 3 (Part 2) includes density, latitude and longitude of each cluster. From Table 3, we can see that all clusters have radius of no longer than two kilometers as expected.

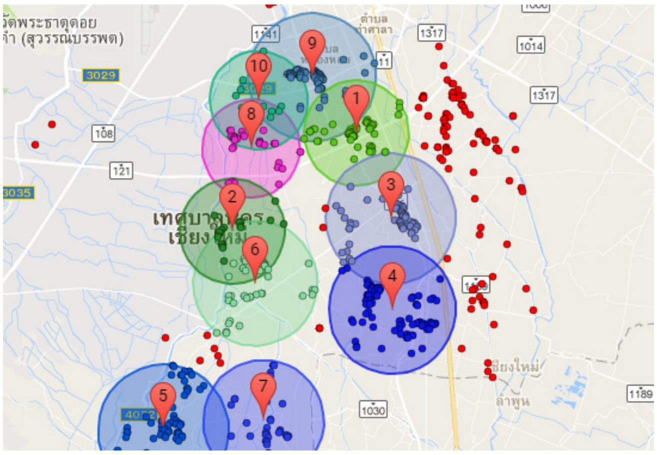


Figure 1. An example of result obtained from the first layer experimented with k = 10 and threshold radius equals two kilometers

Table 3 (Part 1). Ten clusters resulting from the first layer clustering of Saraphi district sorted in descending order with respect to densities

Order	Cluster Number	The Number of Breeding Locations	Radius (km)
1	5	157	1.928486132
2	9	57	1.592775288
3	2	49	1.492037771
4	6	61	1.812016425
5	4	59	1.914995651
6	8	50	1.944217512
7	10	38	1.714531436
8	3	28	1.666085644
9	7	20	1.730876673
10	1	19	1.951709419

Table 3 (Part 2). Ten clusters resulting from the first layer clustering of Saraphi district sorted in descending order with respect to densities

Order	Density (locations/km^2)	Latitude (°)	Longitude (°)
1	13.43744623	18.74171xx	99.014864xx
2	7.15180995	18.646000xx	98.975422xx
3	7.006265201	18.680192xx	99.042180xx
4	5.913651157	18.686581xx	99.026439xx
5	5.121138896	18.710517xx	99.038789xx
6	4.210468574	18.635496xx	98.960685xx
7	4.114743988	18.654033xx	98.998110xx
8	3.210801919	18.66961xx	98.979203xx
9	2.124945878	18.705334xx	99.020074xx
10	1.58771816	18.68293xx	99.061967xx

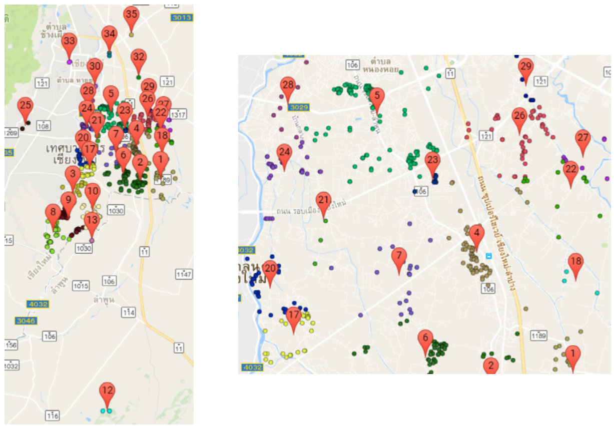
3.2 Experimental Results after the Second Layer of Hierarchical Clustering

According to Figure 1 and Table 3, there are still some locations which do not conform to constraint of the first step of clustering. These locations must be further clustered based on the second layer of clustering and Table 4 and Figure 2 show the obtained result at this step. Note that Table 4 has been separated into two parts, Part 1 and Part 2, where Table 4 (Part 1) includes the number of breeding locations and radius of each cluster while Table 4 (Part 2) includes density, latitude and longitude of each cluster. There are 36 clusters in total discarding the number of locations or households detected in each cluster. Note that there are some clusters with only individual household due to the constraint of threshold radius. This is because any individual detection which is too far from any cluster with a certain density value hardly causes any breeding of common house mosquitoes and thus provide less tendency of diseases with mosquito as a carrier will be breakout. However, such individual location should exactly not be combined into any other cluster and should be left alone unless the original density of the cluster being combined can be improperly biased by the additional individual location.

3.3 Experimental Results with Falsely Alarmed Clusters Filtered Out

An obvious experimental result of falsely alarmed cluster has density equal to 296.0835231 locations per km^2 which is considerably high in a small region while there are only two locations as members in the cluster. This case is geographically displayed in Figure 3 below.

Thus, Table 5 below shows an example that results from filtering out clusters whose radius is shorter than one kilometer and members of smaller than a threshold equal to 50 locations or households. Note that Table 5 has been separated into two parts, Part 1 and Part 2, where Table 5 (Part 1) includes the number of breeding



(a) (b)

Figure 2. The geographical result of 36 clusters obtained from the second layer in Saraphi district, different clusters are colored differently to show members in each cluster: (a) overall view; (b) some part of a close-up view

Table 4 (Part 1). Total of 36 clusters obtained from the second layer clustering of Saraphi district sorted in descending order with respect to densities

Order	Cluster Number	The Number of Breeding Locations	Radius (km)
1	13	3	0.019391214
2	33	2	0.046369556
3	30	2	0.074332944
4	34	2	0.140161872
5	5	157	1.928486132
6	23	9	0.539621461
7	17	26	0.944310606
8	20	27	1.045542098
9	26	41	1.331983776
10	9	57	1.592775288
11	2	49	1.492037771
12	24	31	1.277813301
13	6	61	1.812016425
14	12	3	0.403629313
15	4	59	1.914995651
16	8	50	1.944217512
17	10	38	1.714531436
18	25	2	0.405310298
19	27	8	0.817830894
20	3	28	1.666085644
21	22	20	1.527186919
22	18	3	0.602325933
23	29	5	0.805618737
24	7	20	1.730876673
25	28	7	1.182825674
26	1	19	1.951709419
27	21	6	1.12077177
28	11	1	-
29	14	1	-
30	15	1	-
31	16	1	-
32	19	1	-
33	31	1	-
34	32	1	-
35	35	1	-
36	36	1	-

Table 4 (Part 2). Total of 36 clusters obtained from the second layer clustering of Saraphi district sorted in descending order with respect to densities

Order	Density (locations/km^2)	Latitude (°)	Longitude (°)
1	2539.577032	18.628149xx	98.996839xx
2	296.0835231	18.789015xx	98.975918xx
3	115.2172359	18.766795xx	99.000197xx
4	32.40562058	18.7957xx	99.0134xx
5	13.43744623	18.74171xx	99.014864xx
6	9.838167433	18.727059xx	99.028220xx
7	9.280977936	18.691846xx	98.994715xx
8	7.86196025	18.702366xx	98.98901xx
9	7.355905037	18.737181xx	99.049112xx
10	7.15180995	18.646000xx	98.975422xx
11	7.006265201	18.680192xx	99.042180xx
12	6.04333977	18.728896xx	98.992441xx
13	5.913651157	18.686581xx	99.026439xx
14	5.861462414	18.474367xx	99.011509xx
15	5.121138896	18.710517xx	99.038789xx
16	4.210468574	18.635496xx	98.960685xx
17	4.114743988	18.654033xx	98.998110xx
18	3.875295704	18.7313xx	98.9346xx
19	3.807264852	18.732034xx	99.06434xx
20	3.210801919	18.66961xx	98.979203xx
21	2.729579401	18.724962xx	99.061411xx
22	2.632135597	18.703956xx	99.062727xx
23	2.452228945	18.748377xx	99.05061xx
24	2.124945878	18.705334xx	99.020074xx
25	1.592599892	18.744138xx	98.993398xx
26	1.58771816	18.68293xx	99.061967xx
27	1.520432034	18.717903xx	99.001898xx
28	-	18.9024xx	100.523xx
29	-	18.649932xx	98.170056xx
30	-	18.6787xx	990.13xx
31	-	18.6796xx	9.0453xx
32	-	18.702910xx	98.48200xx
33	-	18.774xx	99.732xx
34	-	18.77517xx	99.040874xx
35	-	18.813311xx	99.034315xx
36	-	17.456928xx	99.013734xx

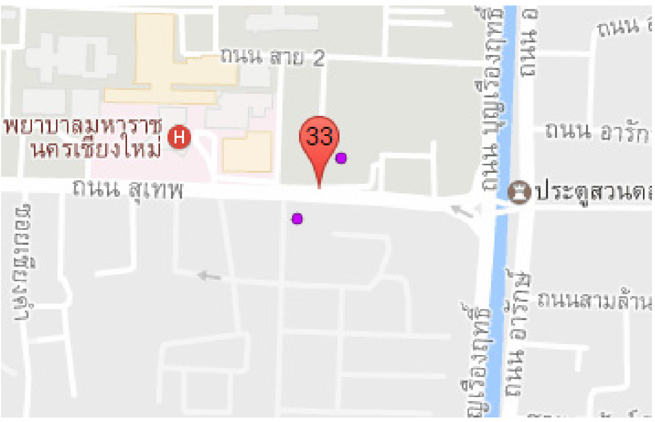


Figure 3. An example of result of falsely alarmed cluster where there are only two households in a small area but its calculated density is considerably high

Table 5 (Part 1). Five clusters in Saraphi district with highest rank densities after falsely alarmed clusters were filtered out

Order	Cluster Number	The Number of Breeding Locations	Radius (km)
1	5	157	1.928486132
2	9	57	1.592775288
3	2	61	1.812016425
4	6	59	1.914995651
5	4	50	1.944217512

Table 5 (Part 2). Five clusters in Saraphi district with highest rank densities after falsely alarmed clusters were filtered out

Order	Density (locations/km^2)	Latitude (°)	Longitude (°)
1	13.43744623	18.74171xx	99.014864xx
2	7.15180995	18.646000xx	98.975422xx
3	5.913651157	18.686581xx	99.026439xx
4	5.121138896	18.710517xx	99.038789xx
5	4.210468574	18.635496xx	98.960685xx

locations and radius of each cluster while Table 5 (Part 2) includes density, latitude and longitude of each cluster.

4 Conclusion

From our experiments, the clustering results showed that our proposed method of clustering or grouping regions of households under a certain given constraint in clustering works very well and the obtained clusters can cover all target locations while the strictly required constraint can still meet.

Acknowledgements

Input data for this research has been provided by Ekkarat Boonchieng, Director of The Center of Excellence in Community Health Informatics (COEiCHI), Chiang Mai University, Thailand. The data was collected in real time on an online application developed by COEiCHI. This work was supported by Thailand Science Research and Innovation under the project IRN62W0007.

References

[1] C.-F. Tsai, C.-T. Tsai, C.-S. Hung, P.-S. Hwang, Data Mining Techniques for Identifying Students at Risk of Failing a Computer Proficiency Test Required for Graduation, *Australasian Journal of Educational Technology*, Vol. 27, No. 3, pp. 481-498, June, 2011.

[2] I. H. Witten, E. Frank, M. A. Hall, *Data Mining: Practical Machine Learning Tools and Techniques*, Morgan Kaufmann, 2011.

[3] L. Parsons, E. Haque, H. Liu, Subspace Clustering for High Dimensional Data: A Review, *ACM SIGKDD Explorations Newsletter*, Vol. 6, No. 1, pp. 90-105, June, 2004.

[4] K. Wagstaff, C. Cardie, S. Rogers, S. Schroedl, Constrained k-Means Clustering with Background Knowledge, *The Eighteenth International Conference on Machine Learning*, Williamstown, MA, USA, 2001, pp. 577-584.

[5] Y. Benkaouz, M. Erradi, B. Freisleben, Work in Progress: k-Nearest Neighbors Techniques for ABAC Policies Clustering, *ACM International Workshop on Attribute based Access Control*, New Orleans, Louisiana, USA, 2016, pp. 72-75.

[6] S. Nutanong, R. Zhang, E. Tanin, L. Kulik, Analysis and Evaluation of V*-kNN: An Efficient Algorithm for Moving kNN Queries, *The International Journal on Very Large Data Bases*, Vol. 19, No. 3, pp. 307-332, June, 2010.

[7] C. Xia, W. Hsu, M. L. Lee, ERkNN: Efficient Reverse k-Nearest Neighbors Retrieval with Local kNN-Distance Estimation, *The 14th ACM International Conference on Information and Knowledge Management*, Bremen, Germany, 2005, pp, 533-540.

[8] J. Macqueen, Some Methods for Classification and Analysis of Multivariate Observations, *The Fifth Berkeley Symposium on Mathematical Statistics and Probability*, Vol. 1, pp. 281-297, 1967.

[9] J. Han, M. Kamber, J. Pei, *Data Mining: Concepts and Techniques*, Elsevier, 2011.

[10] J. Zhang, Q. Yuan, S. Chen, H. Huang, X. Wang, Data Centers Selection for Moving Geo-distributed Big Data to Cloud, *Journal of Internet Technology*, Vol. 20, No. 1, pp. 111-122, January, 2019.

[11] S. Lee, G. Jeon, Mimic Big Data and Low Power Infrastructure-based Small Blood Pressure Measurement for Internet of Things, *Journal of Internet Technology*, Vol. 20, No. 1, pp. 315-322, January, 2019.

[12] W. Boonchieng, E. Boonchieng, W. Tuanrat, C. Khuntichot, K. Duangchaemkarn, Integrative System of virtual Electronic Health Record with Online Community-based Health Determinant Data for Home Care Service: MHealth Development and Usability Test, *2017 IEEE Healthcare Innovations and Point of Care Technologies (HI-POCT)*, Bethesda, MD, 2017, pp. 5-8.

[13] E. Boonchieng, W. Boonchieng, W. Senaratana, J. Singkaew, Development of mHealth for Public Health Information Collection, with GIS, Using Private Cloud: A Case Study of Saraphi District, Chiang Mai, Thailand, *2014 International Computer Science and Engineering Conference (ICSEC)*, Khon Kaen, Thailand, 2014, pp. 350-353.

[14] W. Nadda, W. Boonchieng, E. Boonchieng, Dengue Fever Detection using Long Short-term Memory Neural Network, *2020 17th International Conference on Electrical Engineering/ Electronics, Computer, Telecommunications and Information Technology (ECTI-CON)*, Phuket, Thailand, 2020, pp. 755-758.

Biographies

Varin Chouvatut graduated with B.Eng. (Honours) and M.Eng. in Computer Engineering and got Ph.D. in Electrical and Computer Engineering from King Mongkut's University of Technology Thonburi since 2011. Recently, she is an assistant professor at Chiang Mai University. Her research interests include computer vision, image processing, computer graphics, and data science.

Ekkarat Boonchieng got Ph.D. in Computer Science from Illinois Institute of Technology since 2000. He is currently a Director of Center of Excellence in Community Health Informatics, Chiang Mai University. His research interests include computer graphics, image processing, computer network, data science and biomedical engineering.

Waraporn Boonchieng has completed a Dr. P.H. from Mahidol University, Thailand. She is an Associate Professor in Faculty of Public Health, Chiang Mai University, Thailand. Her research focuses on public health informatics and health promotion.