SPATIO-TEMPORAL CLUSTERING OF DRUG ADDICTS

NOOR AISYAH RASHIF[1], AIDA MELISSA MOHD SHUKRI[2], WAN FAIROS WAN YAACOB[3],
NIK NUR FATIN FATIHAH SAPRI[4], AND MUHAMMAD AMIR ZAHIEN AMRAN[5]

ABSTRACT. This paper mapped spatio-temporal heterogeneity drug addicts' cases in Malaysia within 9 years to understand the geographical area of states and to cluster the hotspot areas of high-risk drug addicts. Spatial maps were used to visualize the drug addicts hotspot areas in 16 states of Malaysia for 9 years (2010-2018). Using Ward linkage hierarchical clustering, this study characterized the drug-addicts risk areas into several clusters. From the spatial map, the most frequent hotspot areas of drug addiction in Malaysia were identified to be Kedah, Pulau Pinang, Selangor, Kelantan, Johor and Perak. The result from hierarchical clustering generated three drug-addicts risk clusters namely the high risk drug addicted (HRDA) states which were Pulau Pinang, Selangor, Kedah, Kelantan, Perak, Johor in cluster 1, the moderate risk drug addicted (MRDA) states which were Terengganu, Pahang and Federal Territory of Kuala Lumpur in cluster 2 and the low risk drug addicted (LRDA) states which were Negeri Sembilan, Perlis, Melaka, Federal Territory of Putrajaya, Federal Territory of Labuan, Sabah and Sarawak in cluster 3. These findings are beneficial for the respective stakeholders to plan actions and intervention needed to be undertaken to reduce the number of drug addicts focusing on the high-risk area of drug addicts.

1. INTRODUCTION

Drug is a substance taken for both therapeutic purpose and abused substance [1]. The problem of drug use and addiction is a global issue [2]. In many years, statistics compiled for the number of drug addicts in Malaysia showed several pattern. In Kelantan, Terengganu, Sabah and Federal Territory of Labuan, the number of drug addicts recorded showed an increasing trend [3]. New Straits Times also highlighted that 87 areas were denoted as areas with the highest risk [4].

Drug addiction problems are also associated with various social and health related problems [5]. Many awareness programs and rehabilitation on the dangers of drugs had been implemented by government agencies or private parties, but the same problem remains as it is difficult to combat [6]. Several studies have been conducted to understand and interpret the drugs addiction patterns and behaviour. Among the popular techniques used are Geographic information system (GIS) technology [7] and spatial analysis. Besides that, researchers also used the method of clustering. Toriman et al. [7] used clustering to identify the hot spot area and to classify it into high drug addicted cluster in the state of Terengganu, Malaysia. Thus, this study uses the Wards linkage hierarchical clustering technique to characterize the drug-addicts risk area into several potential clusters.

2. METHODS/MATERIALS

2.1. **Data Description.** This study used the secondary data source obtained from National Anti-Drugs Agency (NADA). This study is a retrospective cohort study that used the record of drug addiction covering all states in Malaysia for a period of 9-year from 2010 to 2018. Table 1 presents the description of the data used in this research.

TABLE 1. Sources of data

Data	Period	Measurement
Yearly number of drug addicts in Malaysia by state	9-year period (2010 − 2018)	Number of drug addicts
Longitude	16 states in Malaysia	Degree, Minutes, Seconds
Latitude	16 states in Malaysia	Degree, Minutes, Seconds

2.2. **Methods of Data Analysis.**

2.2.1. *Spatial and Temporal analysis.* The spatial map was plotted for each state with different colour tones using R software to display the hotspot area of drug addiction in Malaysia. Temporal analysis was also used throughout the 9-year period, which is a detailed timeline analysis that could relate towards the drug addict pattern and potentially lead to other sources of evidence [8] across the time.

2.2.2. *Ward Hierarchical Linkage Clustering.* In this study, Ward hierarchical linkage agglomerative clustering was used to the normalized datasets by means of the ward method. On the basis of single Euclidean measures similarity, the object was classified, and the classification of the object was illustrated in a dendrogram. The Ward's method of distance determination is a hierarchical agglomerative approach that treats each observation as a single cluster and pair of clusters which emerge as one that moves up the hierarchical cluster of the object into a bigger cluster. This algorithm was used in defining the compact clusters and finding a partition with small sum of squares. In Ward Hierarchical Clustering, this method measures the distance between A and B by how much the sum of squared will increase when two points are merged:

$$\Delta(A, B) = \sum_{i \in A \cup B} ||\vec{x}_i - \vec{m}_{A \cup B}||^2 - \sum_{i \in A} ||\vec{x}_i - \vec{m}_A||^2 - \sum_{i \in B} ||\vec{x}_i - \vec{m}_B||^2$$
$$= \frac{n_A n_B}{n_A + n_B} ||\vec{m}_A - \vec{m}_B||^2,$$

where $\vec{m}_j$ is the center of cluster j, and $\vec{n}_j$ is the number of points in the cluster j. Δ is called the merging cost of combining the clusters A and B. Ward's method does not rely on a random starting guess, so its answer is unique [9].

3. RESULTS AND DISCUSSIONS

3.1. **Descriptive Analysis.**

3.1.1. *Descriptive Analysis of Drug Addicts by States.* The distribution on the number of drug addicted person as according to each state is presented by the boxplot in Fig. 1. The boxplot shows a significant difference in the distribution of drug addicted person among the states in Malaysia. Pulau Pinang spotted the

highest ranking in the average number of drug addicts compared to the other 14 states. It is then followed closely by Kedah and Perak. The Federal Territory of Putrajaya had the lowest average number of drug addicts, since Putrajaya recorded the lowest population among the 16 states.
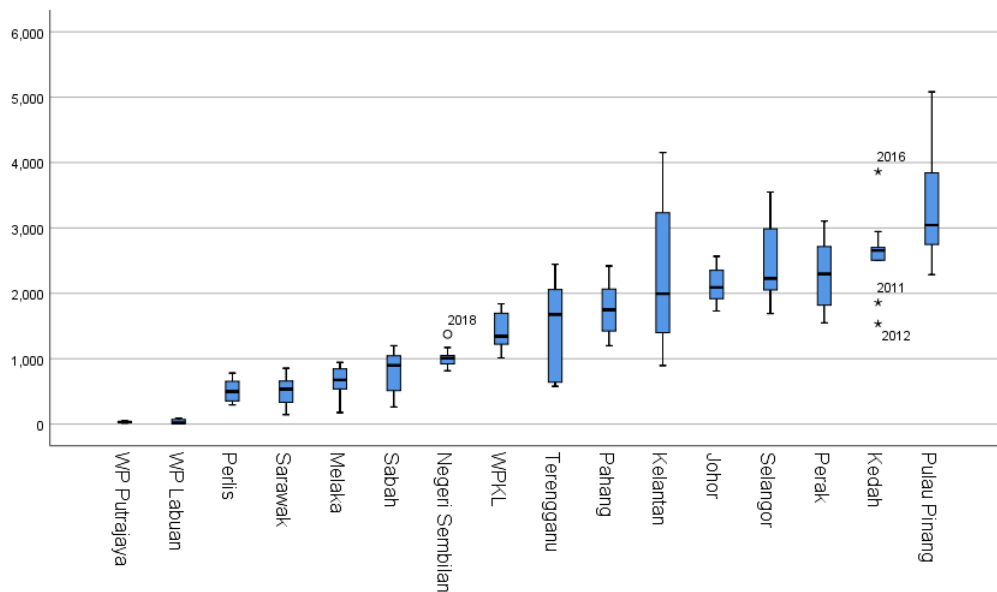

FIGURE 1. The Boxplot of the Number of Drug Addict by States

3.2. **Spatial and Temporal Analysis.**

3.2.1. *Spatial mapping of Drug Addicts in Peninsular Malaysia.* The maps demonstrated the results of analysis to visualize the differences in the number of drug addicts across Peninsular Malaysia and East of Malaysia, the island of Borneo. Spatial distribution of drug addicts in Peninsular Malaysia by year is described by a series of features in Fig. 2 of A) to I). The spatial distributions showed a few shades of red. The darkest shades of red indicated that the area was in the category of very high count of drug addicts whereas the brightest red indicated very low count of drug addicts.

The spatial mapping showed significant changes throughout the 9 years. The first map in Fig 2(A) is the mapping for 2010 which showed that 3 out of 13 states had the darkest shades of red.
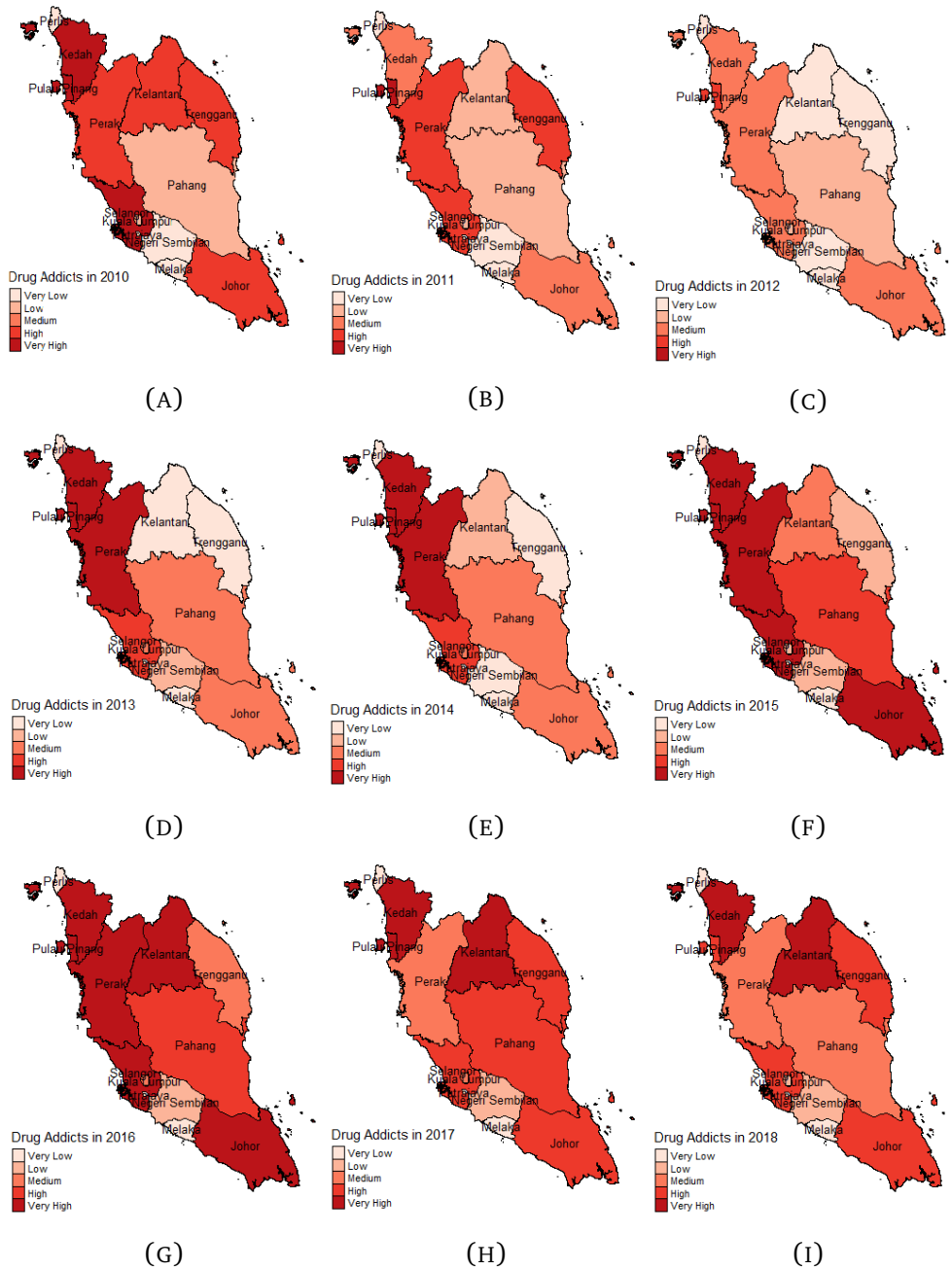
FIGURE 2. Spatial Mapping of Drug Addicts in Peninsular Malaysia in A) 2010, B) 2011, C) 2012, D) 2013, E) 2014, F) 2015, G) 2016, H) 2017, I) 2018

The darkest shades of red keep changing. The number of drug addicts in Kedah was seen to change at first from dark shades (high) of red to the brightest shades, then the colour changed again to the darkest shades (very high). Kedah in 2018 still became one of the hotspot areas in Peninsular Malaysia. Other than Kedah, the pattern in the number of drug addicts in Kelantan was also seen to change from a high number of drug addicts to very high number of drug addicts within the 9 years. Among the unique significant explanation for Kedah and Kelantan is the geographical position of both states that are very close to the border of Thailand. This could motivate drug trafficking activities such as access to the source supply and distribution. Therefore, both states need to be monitored more often.

3.2.2. *Spatial mapping of Drug Addicts in East Malaysia.* The spatial mapping of drug addicts in East Malaysia is plotted in Fig. 3. East Malaysia comprises of Sabah, Sarawak and Federal Territory of Labuan. The map show both Federal Territory of Labuan and Sarawak indicated a very low state of drug addiction risk. The distribution was constant for 9 years under study. However, for the state of Sabah, the number of drug addicts showed an increasing pattern from very low to low risk area of drug addicts.

3.2.3. *Temporal Analysis of Drug Addicts.* The distribution trend of drug addicts for every state in Malaysia is shown in Fig. 4 (A-P). Selangor, Terengganu, Kedah, and Pulau Pinang showed an increasing trend for 9 years while Negeri Sembilan, Kelantan, Sabah, Sarawak, Perak, Melaka and Johor displayed increasing and decreasing patterns from 2010 until 2016, but showed a sudden slight increase in the years 2017 and 2018.

3.2.4. *Clustering Analysis of Drug Addicts by States.* The dendrogram obtained from Wards Hierarchical Linkage clustering is displayed in Fig. 5. The first cluster indicates the high risk of drug addict areas which were Pulau Pinang, Selangor, Kedah, Kelantan, Perak and Johor. The second cluster indicates medium risk drug addict areas which were Terengganu, Pahang and Federal Territory of Kuala Lumpur. Lastly East Malaysia of Sabah, Sarawak and Federal Territory of Labuan, Negeri Sembilan, Perlis, Melaka and Federal Territory of Putrajaya were
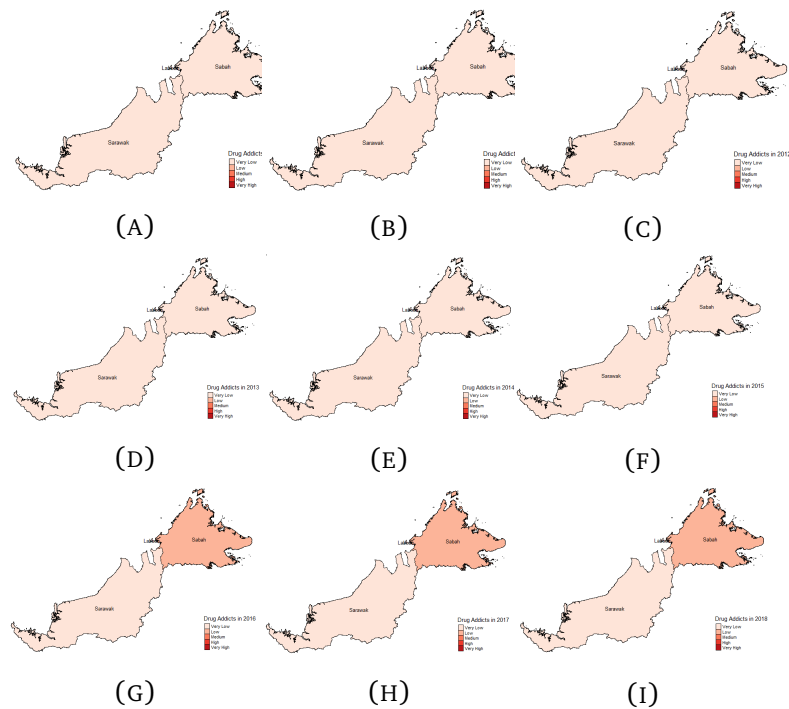
FIGURE 3. Spatial Mapping of Drug Addicts in East Malaysia for A) 2010, B) 2011, C) 2012, D) 2013, E) 2014, F) 2015, G) 2016, H) 2017, I) 2018

in the third cluster of low risk drug addict areas. The list of cluster membership by states that represent each cluster is described in Table 2.

4. CONCLUSION

This paper aims to cluster and identify the hotspot areas of drug addicts in every state of Malaysia using spatial analysis within 9-year period of 2010 to 2018. The spatial maps obtained indicated that the distribution of drug addicts in every state was contributed by their geographical area. For instance, the location of the states such as Kedah and Kelantan which are near to the border of Thailand enables an easier supply of drugs activities of supply and distribution to occur which leads those state in the high risk areas. Selangor, Negeri Sembilan, Terengganu, Perak, Johor, and Pulau Pinang that have industrial areas also contributed inti high risk areas of drug addicts. Sabah and Sarawak, which
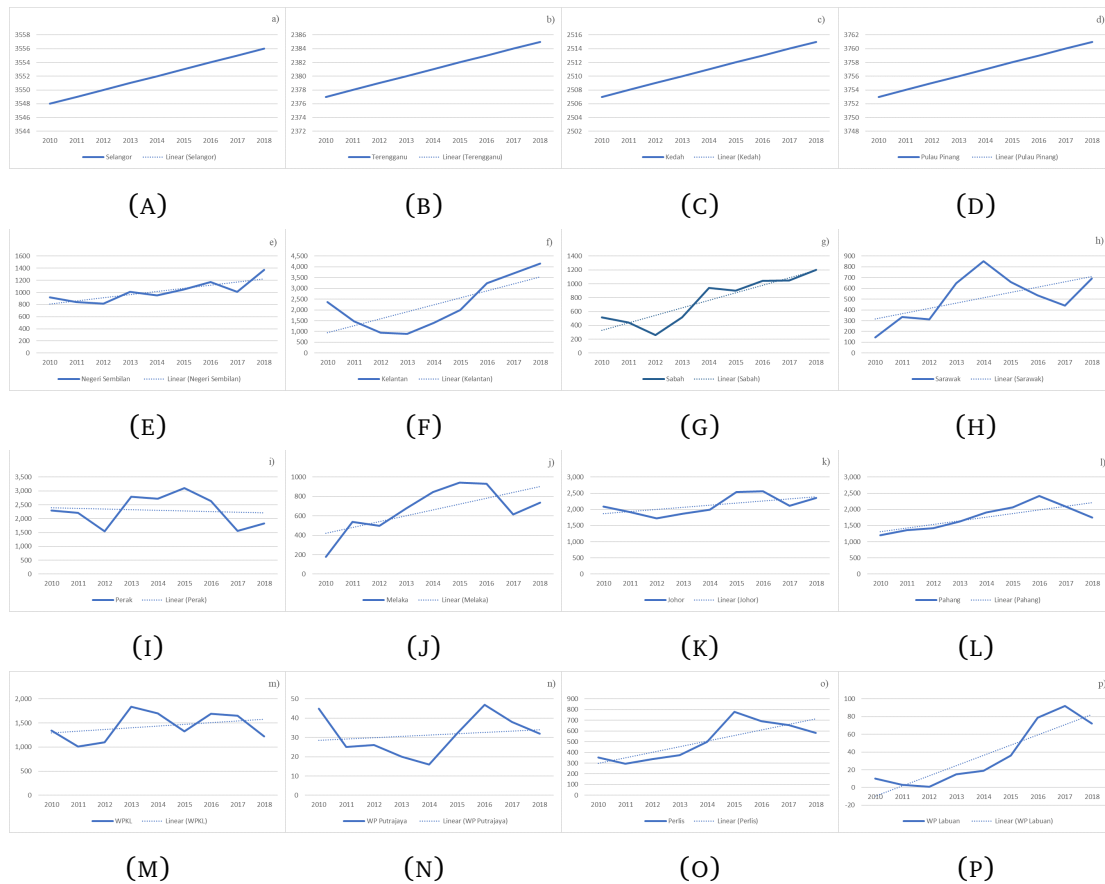
FIGURE 4. Temporal distribution of drug addicts by states A) Selangor, B) Terengganu, C) Kedah, D) Pulau Pinang, E) Negeri Sembilan, F) Kelantan, G) Sabah, H) Sarawak, I) Perak, J) Melaka, K) Johor, L) Pahang, M) Federal Territory of Kuala Lumpur, N) Federal Territory of Putrajaya, O) Perlis, P) Federal Territory of Labuan.

are located in the West of Malaysia and far from Peninsular Malaysia, showed a lower number of drug addicts. The maps also indicated that small area of states such as Perlis, Melaka, Federal Territory of Putrajaya, Federal Territory of Labuan contained a lower number of drug addicts.

This study also clusters the drug addict's incidence according to state using cluster analysis. Based on the results obtained, the finding came up with three levels of cluster which were high risk drug addict area (HRDA), medium risk
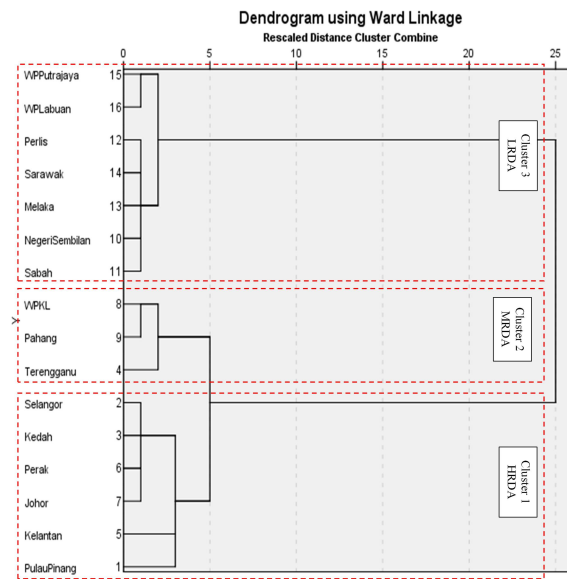
FIGURE 5. Hierarchical dendrogram for clustering drug addicts by states

TABLE 2. Cluster Membership by State

Cluster 1 HRDA	Cluster 2 MRDA	Cluster 3 LRDA
Pulau Pinang	Terengganu	Negeri Sembilan
Selangor	Pahang	Sabah
Kedah	Federal Territory of Kuala Lumpur	Perlis
Kelantan		Melaka
Perak		Sarawak
Johor		Federal Territory of Putrajaya
		Federal Territory of Labuan

drug addict area (MRDA) and low risk drug addict area (LRDA) as shown in the dendrogram plot. The HRDA cluster includes Pulau Pinang, Selangor, Kedah, Kelantan, Perak and Johor, while the second cluster of MRDA consists of Terengganu, Federal Territory of Kuala Lumpur and Pahang and the rest of the states are in the LRDA of cluster which recorded low number of drug addicts.

REFERENCES

[1] M.N. ONGWAE: *A study of the causes and effects of drug and substance abuse among students in selected secondary schools in Starehe Sub County, Nairobi County*, Doctoral dissertation, University of Nairobi.

[2] A. RAZALI, Z. MADON: *Issues and challenges of drug addiction among students in Malaysia*, Advances in Social Sciences Research Journal, **3**(8) (2016), 231–238.

[3] DEPARTMENT OF STATISTICS MALAYSIA: Available from: https://www.dosm.gov.my, (2018).

[4] H.B. NURUL: *More than 1,700 secondary school students test positive for drug use. New Straits Times*, Available from https://www.nst.com.my/news/crime-courts/2018/12/444639/more-1700-secondary-school-students-test-positive-drug-use (28th December 2018).

[5] M. MOHAMAD, M. MOHAMMAD, I. MAMAT, M. MAMAT: *Modelling Positive Development, Life Satisfaction and Problem Behaviour among Youths in Malaysia*, World Applied Sciences Journal, **32**(2) (2014), 231–238.

[6] H.H MOHD NOOR: *Kelantan has highest number of new drug abusers*, New Straits Time. https://www.nst.com.my/news/nation/2018/06/384225/kelantan-has-highest-number-new-drug-abusers-nada (26th June 2008).

[7] M.E. TORIMAN, S.N. ABDULLAH, I.A. AZIZAN, M.K. KAMARUDIN, R. UMAR, N. MOHAMAD : *Spatial and temporal assessment on drug addiction using multivariate analysis and GIS*, Malaysian Journal of Analytical Sciences, **19**(6) (2015), 1361–1373.

[8] E. CASEY: *Handbook of digital forensics and investigation*, Academic Press, 2009.

[9] C. SHALIZI: *Distances between Clustering. Hierarchical Clustering Data Mining*, (2009), 36–50.

[1,2,3]FACULTY OF COMPUTER AND MATHEMATICAL SCIENCES, UNIVERSITI TEKNOLOGI MARA CAWANGAN KELANTAN, KAMPUS KOTA BHARU, LEMBAH SIREH,15050 KOTA BHARU, KELANTAN, MALAYSIA

Email address: [1]aisyahrashif@gmail.com

Email address: [2]aidamelissashukri@gmail.com

Email address: [3]wnfairos@uitm.edu.my

[4]FACULTY OF COMPUTER AND MATHEMATICAL SCIENCES, UNIVERSITI TEKNOLOGI MARA, 45400 SHAH ALAM, SELANGOR, MALAYSIA.

Email address: fatinfatihah26@yahoo.com

[5]NATIONAL ANTI-DRUGS AGENCY, STATE DRUG AGENCY, JALAN MAKTAB PERGURUAN ISLAM, 43000 KAJANG, SELANGOR, MALAYSIA

Email address: amir.zahien@adk.gov.my