Comparison Between K-Means and Fuzzy C-Means Clustering in Network Traffic Activities

Purnawansyah[1], Haviluddin[2(✉)], Achmad Fanany Onnilita Gafar[3], and Imam Tahyudin[4]

[1] Faculty of Computer Science, Universitas Muslim Indonesia, Makassar, Indonesia
[2] Faculty of Computer Science and Information Technology, Mulawarman University, Samarinda, Indonesia
haviluddin@gmail.com
[3] State Polytechnic of Samarinda, Samarinda, Indonesia
[4] Information System Department, STMIK AMIKOM, Purwokerto, Indonesia

Abstract. A network traffic utilization in order to support teaching and learning activities are an essential part. Therefore, the network traffic management usage is requirements. In this study, analysis and clustering network traffic usage by using K-Means and Fuzzy C-Means (FCM) methods have been implemented. Then, both of method were used Euclidean Distance (ED) in order to get better results clusters. The results showed that the FCM method has been able to perform clustering in network traffic.

Keywords: Network traffic · K-Means · Fuzzy C-Means · Clustering

1 Introduction

A bandwidth usage monitoring management in a network traffic at university is indispensable. Therefore, a recording and analysis of bandwidth usage by network administrators is very beneficial. It aims to make use of the bandwidth can be well controlled, stable access, and gives users convenient access. Therefore, mapping or cluster of bandwidth usage in order to support the network administrator performance analysis is required.

In this study, the cluster method based on intelligence algorithm are proposed in bandwidth usage, such as Self-Organizing Maps (SOM), K-Means, Fuzzy C-Means, etc. For that reason, these algorithms have been widely used by researchers in order to solve cluster data problems in a variety of fields, including economics [5], supply chain [2], engineering [6], hydrology [1,4], internet and social media [3,8], pattern recognition [7], and so forth. Many research results have shown that the algorithms are able to provide accurate information in solving the clustering problem.

J. Xu et al. (eds.), *Proceedings of the Eleventh International Conference on Management Science and Engineering Management*, Lecture Notes on Multidisciplinary Industrial Engineering, DOI 10.1007/978-3-319-59280-0_24

Afterward, two intelligence algorithms is K-Means and Fuzzy C-Means (FCM) for cluster bandwidth usage data has been implemented. The purpose of this study are compared the feasibility of two clustering methods and how it works in the real world problems particularly subject on network traffic. Thus, the expected results of network traffic clustering are coordination schemes that support resource management [8]. Furthermore, this paper will apply two models, namely K-Means and Fuzzy C-Means (FCM) that have been developed and compared in order to cluster the network traffic usage. Section 2 describes the architectures of K-Means and FCM clustering models are proposed. Section 3 describes the analysis and discussion of the results. Finally, conclusions are summarized in Sect. 4.

2 Research Method

In this section, a brief information of K-Means and FCM models are presented.

2.1 Principle of K-Means Method

K-means clustering is an unsupervised learning classified. This algorithm is based on the determination of the distance between the centroid and the training data. Then, the number of cluster centroid based on the number desired. Meanwhile, the initialization centroid randomly generated by considering the data training. In other words, the centroid should be in the training data space. Then, a couple of training data is from the attributes of the data patterns to be planned. In each iteration, the distance of each training data with each centroid to be calculated. It means that any training data will have a centroid distance. At the same time, members of the cluster indicated by the smallest distance from the corresponding centroid. Then, a new centroid value is calculated based on the average value of each member of the cluster. If the cluster member does not change then the iteration is stopped, Fig. 1.

The following data clustering techniques using the K-Means algorithm as follows.

- Determine the number of clusters K.
- The Initialization of K cluster centers can be done randomly and used as initial cluster centers.
- Allocate all data/objects to the nearest cluster. The proximity of the two objects is determined by the distance of the object. To calculate the distance of all the data to the cluster center point using the Euclidean distance theory formulated as follows:

$$T_{(x,y)} = \sqrt{(T_{1x} - T_{1y})^2 + (T_{2x} - T_{2y})^2 + \cdots + (T_{kx} - T_{ky})^2}, \quad (1)$$

where, $T_{(x,y)}$ is distance data of x to the cluster center of y; T_{kx} is i-data on attribute data of k; T_{ky} is the center point of j on the attribute of k.

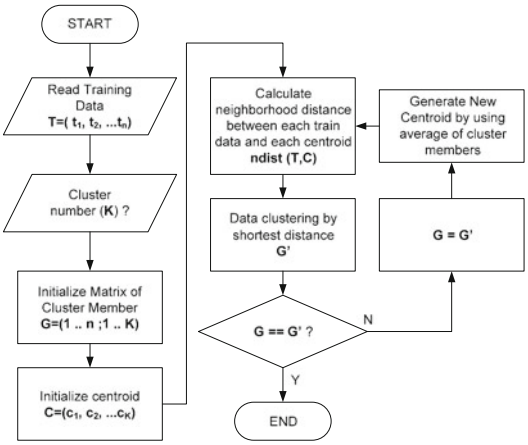


Fig. 1. K-Means algorithm

- Recalculate the center of the cluster with the new cluster membership. This is computed by determining the centroid/center cluster.
- Assign each object to put on the new cluster center, if the center of the cluster changed, back to 3, otherwise clustering is complete.
- Analyze the results in the clustering process.

2.2 Principle of Fuzzy C-Means Method

Other advanced techniques clustering in machine learning model come from the Dunn in 1973, and improved by Bezdek in 1981, called Fuzzy C-Means (FCM) clustering was developed by Dunn in 1973, and improved by Bezdek in 1981. In principle, FCM clustering process is based on a partition of a set of data into a number of clusters with minimum similarity between different clusters [2]. Since the introduction of the fuzzy set theory in 1965 by Zadeh, it has been applied in a variety of fields. FCM is a flexible fuzzy partition that an improvement of common C-Means algorithm [5]. At FCM, each feature vectors valued between [0–1] by using the membership function, because FCM is based on the criteria of distances numbers between clusters. In other words, FCM clustering is based on Point-Prototype Clustering Model with output centroid most optimal partition. Where, partition optimization centroid is obtained by minimizing the objective function. The formula is given by:

$$J_{\mathrm{FCM}}(U,V) = \sum_{j=1}^{N}\sum_{i=1}^{C}(u_{ij})^{q}(d_{ji})^{2},$$

where
$U =$ Fuzzy datasets K-partition;
$V =$ Set of prototype centroid;
$V = \{v_1, v_2, \cdots, v_C\} \subset R^P$;

$$(d_{ji})^2 = \|x_j - v_i\|^2 = \sqrt{\left(x_{j(\text{row})} - v_{i(\text{row})}\right)^2 + \left(x_{j(\text{col})} - v_{i(\text{col})}\right)^2}. \quad (2)$$

Euclidean Distance between x_j and v_i;
$X = \{x_1, x_2, \cdots, x_n\} \subset R^P$;
$v_i =$ Centroid cluster to i;
$u_{ij} =$ Membership level x_j in cluster to i;
$N =$ Total data;
$C =$ Total cluster;
$q =$ Fuzzifier parameter, $q > 1$.

Afterward, all of the objects in each cluster has a certain degree of proximity or similarity. Meanwhile, the FCM processes consists of five stages. First, to determine the cluster is set become a center cluster location marker on average for each cluster. Second, calculate the distance between feature vector (X) and the centroid vector $(V)[X \to V]$. In this experiment, Euclidean Distance (ED) was implemented. Third, Calculate membership level. Fourth, Calculate new centroid. Lastly, recalculated the new centroid, if criterion between [0–1] is reached then stop the iteration. In this study, the FCM Clustering algorithm is as follows:

Step 1. Initialization Vector centroid, v_i (prototypes).
Step 2. Calculate the distance between feature vector (X) and the centroid vector $(V)[X \to V]$. Feature vectors with the closest distance to one of the centroid vectors then expressed as a cluster member.
Step 3. Calculate membership level of all feature vectors in all clusters by using the formula:

$$u_{ij} = \frac{1}{\sum_{k=1}^{K} \left[\frac{(d_{ji})^2}{(d_{jk})^2}\right]^{1/(q-1)}} = \frac{\left[\frac{1}{(d_{ji})^2}\right]^{1/(q-1)}}{\sum_{k=1}^{K} \left[\frac{1}{(d_{jk})^2}\right]^{1/(q-1)}}. \quad (3)$$

Step 4. Calculate new centroid using Eq. (4).

$$\hat{V}_i = \frac{\sum_{j=1}^{N} (u_{ij})^q X_j}{\sum_{j=1}^{N} (u_{ij})^q}. \quad (4)$$

Step 5. Recalculate the step 4, $u_{ij} \to \hat{u}_{ij}$ If, $\max_{ij} |u_{ij} - \hat{u}_{ij}| < \varepsilon$, where ε is termination criteria between 0 and 1. Then, the iteration process is stopped. If not go back to step 5. The FCM algorithm can be seen in Fig. 2.

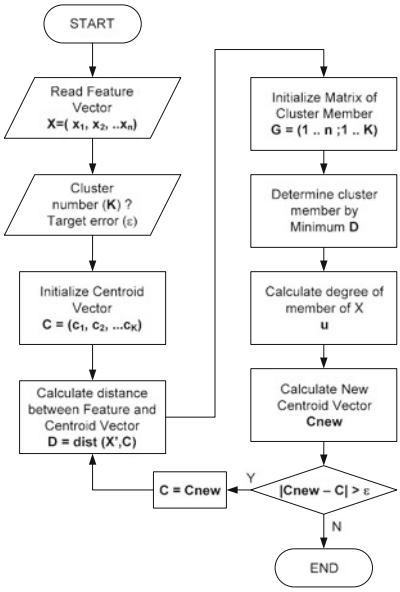


Fig. 2. Fuzzy C-means algorithm

(1) Initialization Centroid
Centroid is used as a central member of the cluster that expected to be an average of all the cluster members. If k cluster is formed, the number k centroid is necessary. There are various ways of doing initialization centroid, such as the random generation and the average spread. For example, interval data is [mindata$\cdots$maxdata]. Centroid random initialization is done with formula $C = \text{mindata} + (\text{maxdata} - \text{mindata}) \times \text{rand}()$. In this study, the initialization centroid for K-Means and FCM are used the average spread then the formula as follows:

$$\begin{aligned}
&\text{cluster number} \;=\; k\\
&s = (\text{maxdata} - \text{mindata})/k\\
&m = \text{mean}(\text{data})\\
&C = [(m - s^*(k)) \;\; (\cdots) \;\; (m - s^*(1)) \;\; (m) \;\; (m + s^*(1)) \;\; (\cdots) \;\; (m + s^*(k))]
\end{aligned}$$

(2) Calculated Distance Neighborhoods
Neighborhoods distance is the distance between a centroid to each data that expressed by Eq. (2). According to [9] "an important step in most clustering is to select a distance measure, which will determine how the similarity of two elements is calculated". Thus, there are some varieties of distance function in

clustering, including Euclidean distance (ED), Manhattan distance, Mahalanobis distance, and Hamming distance. In this study, K-Mean and FCM clustering using ED distance.

(3) Generated New Centroid

In this study, generating new centroid both of K-Means and FCM are based on the average all members cluster values. As an illustrated, for training data is as follows.

$$T = \begin{bmatrix} X \\ Y \end{bmatrix} = \begin{bmatrix} x_1 & x_2 & x_3 & x_4 & x_5 & x_5 \\ y_1 & y_2 & y_3 & y_4 & y_5 & y_6 \end{bmatrix}.$$

Next, the clustering results in an iteration is as follows.

$$G = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 1 \end{bmatrix}.$$

Then, the new centroid is as follows.

$$c_1 = \frac{x_1 + x_3 + x_4}{y_1 + y_3 + y_4} \quad c_2 = \frac{x_2 + x_5 + x_6}{y_2 + y_5 + y_6}.$$

Meanwhile, in FCM, generating new centroid by using the COA (Center of Area) formula, Eq. (4).

2.3 Datasets

In this study, 152 days (from January–May 2016) daily network traffic usage of four client datasets from ICT unit were captured. Then, the data are analyzed by using MATLAB R2013b. The real dataset can be seen in Table 1 (Fig. 3).

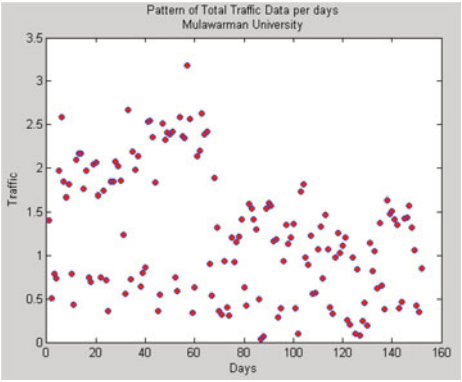


Fig. 3. Plot of network traffic per days

Table 1. The real of daily network traffic data (January–May 2016)

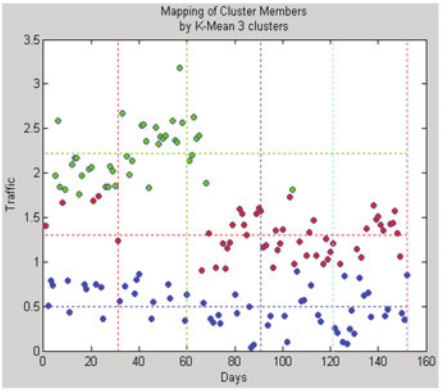
Rectorate	Forestry	Science	Economic	No	Rectorate	Forestry	Science	Economic
0.341	0.391	0.343	0.332	77	0.39	0.365	0.402	0
0.08	0.195	0.061	0.174	78	0.392	0.271	0.358	0.2
0.069	0.468	0.042	0.212	79	0.385	0.32	0.483	0.225
0.097	0.586	0.048	0.007	80	0.08	0.311	0.237	0.004
0.564	0.687	0.352	0.362	81	0.107	0.205	0.069	0.043
0.429	1	0.68	0.476	82	0.415	0.325	0.508	0.348
0.411	0.705	0.467	0.259	83	0.436	0.348	0.485	0.275
0.423	0.569	0.465	0.212	84	0.433	0.312	0.384	0.291
0.379	0.623	0.364	0.447	85	0.331	0.409	0.358	0.203
0.078	0.496	0.044	0.167	86	0.099	0.28	0.115	0.001
0.057	0.345	0.028	0.007	87	0.04	0	0	0.001
0.659	0.583	0.444	0.41	88	0.065	0	0	0.001
0.642	0.724	0.414	0.384	89	0.401	0.384	0.441	0.319
0.471	0.663	0.48	0.556	90	0.412	0.396	0.473	0.324
0.421	0.597	0.295	0.452	91	0.483	0.393	0.413	0.282
0.354	0.45	0.394	0.771	92	0.441	0.501	0	0.227
0.043	0.372	0.119	0.21	93	0.411	0.308	0.09	0.375
0.059	0.269	0.054	0.313	94	0.081	0.131	0.065	0.007
0.507	0.569	0.506	0.46	95	0.045	0.188	0.039	0.116
0.499	0.51	0.445	0.612	96	0.387	0.182	0.166	0.199
...	...	...	...	...	...	...	...	...
...	...	...	...	...	...	...	...	...
0.062	0.213	0.034	0	150	0.06	0.257	0.056	0.055
0.386	0.364	0.457	0.001	151	0.082	0.228	0.014	0.029
0.373	0.289	0.264	0	152	0.362	0.259	0	0.232

3 Results and Analysis

This section presents the empirical work and compares the experimental results of the K-Means and FCM algorithms on highest average usage network traffic problems. The performances are measured by the objective function value given by Eq. (1).

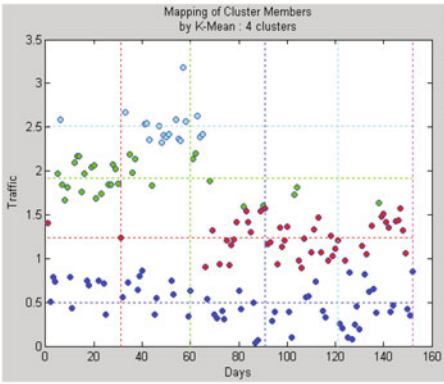
3.1 Analysis of K-Means

In this experiment, the scheme used is to classify the training data into 3, 4, and 5 cluster grouping patterns in order to observe the good cluster. The average data value is used as one of the centroid values. Then, for another centroid value is determined randomly in the space of data training. The results of K-Means are shown in Fig. 4.



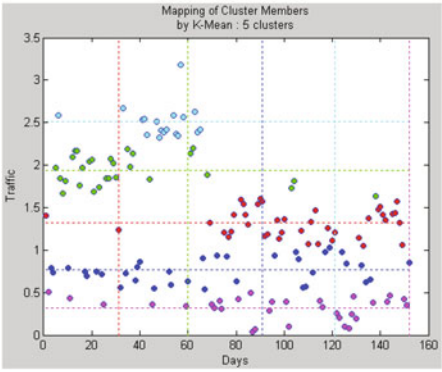
Cluster number/Algorithm : 3/K-mean
Iteration max. : 8
Average of the highest data traffic : 2.4076
Highest number of cluster members : 10
in months : 2

No.	Day	Rektorat	Fahutan	FMIPA	Fekon	Total
1	33	0.4262	0.7720	0.4697	1.0000	2.6679
2	35	0.5053	0.6652	0.3613	0.6528	2.1847
3	36	0.3931	0.8125	0.5078	0.2632	1.9766
4	37	0.4018	0.7949	0.5059	0.4293	2.1319
5	41	0.4310	0.8001	0.6145	0.6831	2.5286
6	42	0.5371	0.8739	0.5657	0.5641	2.5407
7	43	0.4459	0.7911	0.5324	0.5878	2.3572
8	44	0.3209	0.5986	0.4605	0.4550	1.8350
9	47	0.4180	0.8140	0.6014	0.6816	2.5150
10	48	0.4072	0.9090	0.5361	0.4774	2.3296
11	49	0.3915	0.8273	0.5724	0.6154	2.4066
12	50	0.4167	0.7875	0.6895	0.4979	2.3916
13	51	0.3617	0.7700	0.5298	0.7608	2.4223
14	54	0.4249	0.8004	0.6672	0.6901	2.5826
15	55	0.5554	0.7653	0.5491	0.4991	2.3689
16	56	0.5460	0.8762	0.5908	0.3332	2.3462
17	57	0.5833	0.7647	0.9250	0.9108	3.1839
18	58	0.4497	0.9186	0.6768	0.5218	2.5669
Average		0.4453	0.7967	0.5753	0.5902	2.4076



Cluster number/Algorithm : 4/K-mean
Iteration max. : 4
Average of the highest data traffic : 2.5149
Highest number of cluster members : 14
in months : 2

No.	Day	Rektorat	Fahutan	FMIPA	Fekon	Total
1	33	0.4262	0.7720	0.4697	1.0000	2.6679
2	41	0.4310	0.8001	0.6145	0.6831	2.5286
3	42	0.5371	0.8739	0.5657	0.5641	2.5407
4	43	0.4459	0.7911	0.5324	0.5878	2.3572
5	47	0.4180	0.8140	0.6014	0.6816	2.5150
6	48	0.4072	0.9090	0.5361	0.4774	2.3296
7	49	0.3915	0.8273	0.5724	0.6154	2.4066
8	50	0.4167	0.7875	0.6895	0.4979	2.3916
9	51	0.3617	0.7700	0.5298	0.7608	2.4223
10	54	0.4249	0.8004	0.6672	0.6901	2.5826
11	55	0.5554	0.7653	0.5491	0.4991	2.3689
12	56	0.5460	0.8762	0.5908	0.3332	2.3462
13	57	0.5833	0.7647	0.9250	0.9108	3.1839
14	58	0.4497	0.9186	0.6768	0.5218	2.5669
Average		0.4568	0.8193	0.6086	0.6302	2.5149



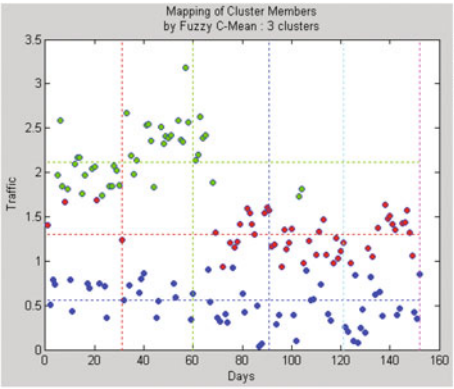
Cluster number/Algorithm : 5/K-mean
Iteration max. : 8
Average of the highest data traffic : 2.5149
Highest number of cluster members : 14
in months : 2

No.	Day	Rektorat	Fahutan	FMIPA	Fekon	Total
1	33	0.4262	0.7720	0.4697	1.0000	2.6679
2	41	0.4310	0.8001	0.6145	0.6831	2.5286
3	42	0.5371	0.8739	0.5657	0.5641	2.5407
4	43	0.4459	0.7911	0.5324	0.5878	2.3572
5	47	0.4180	0.8140	0.6014	0.6816	2.5150
6	48	0.4072	0.9090	0.5361	0.4774	2.3296
7	49	0.3915	0.8273	0.5724	0.6154	2.4066
8	50	0.4167	0.7875	0.6895	0.4979	2.3916
9	51	0.3617	0.7700	0.5298	0.7608	2.4223
10	54	0.4249	0.8004	0.6672	0.6901	2.5826
11	55	0.5554	0.7653	0.5491	0.4991	2.3689
12	56	0.5460	0.8762	0.5908	0.3332	2.3462
13	57	0.5833	0.7647	0.9250	0.9108	3.1839
14	58	0.4497	0.9186	0.6768	0.5218	2.5669
Average		0.4568	0.8193	0.6086	0.6302	2.5149

Fig. 4. Plot of clustering network traffic results using k-means

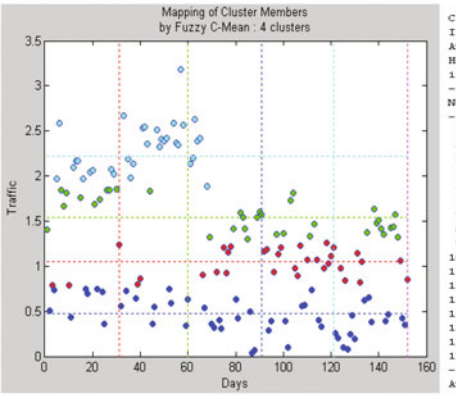
3.2 Analysis of Fuzzy C-Means

In this experiment, the scheme used is to classify the training data into 3, 4, and 5 cluster grouping patterns in order to observe the good cluster. The average data value is used as one of the centroid values. Then, for another centroid value is determined randomly in the space of data training. In this test, the FCM



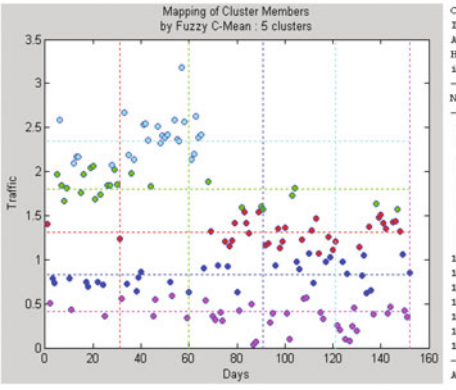
Cluster number/Algorithm : 3/Fuzzy C-mean
Iteration max. : 9
Average of the highest data traffic : 2.4076
Highest number of cluster members : 16
in months : 2

No.	Day	Rektorat	Fahutan	FMIPA	Fekon	Total
1	33	0.4262	0.7720	0.4697	1.0000	2.6679
2	35	0.5053	0.6652	0.3613	0.6528	2.1847
3	36	0.3931	0.8125	0.5078	0.2632	1.9766
4	37	0.4018	0.7949	0.5059	0.4293	2.1319
5	41	0.4310	0.8001	0.6145	0.6831	2.5286
6	42	0.5371	0.8739	0.5657	0.5641	2.5407
7	43	0.4459	0.7911	0.5324	0.5878	2.3572
8	44	0.3209	0.5986	0.4605	0.4550	1.8350
9	47	0.4180	0.8140	0.6014	0.6816	2.5150
10	48	0.4072	0.9090	0.5361	0.4774	2.3296
11	49	0.3915	0.8273	0.5724	0.6154	2.4066
12	50	0.4167	0.7875	0.6895	0.4979	2.3916
13	51	0.3617	0.7700	0.5298	0.7608	2.4223
14	54	0.4249	0.8004	0.6672	0.6901	2.5826
15	55	0.5554	0.7653	0.5491	0.4991	2.3689
16	56	0.5460	0.8762	0.5908	0.3332	2.3462
17	57	0.5833	0.7647	0.9250	0.9108	3.1839
18	58	0.4497	0.9186	0.6768	0.5218	2.5669
Average		0.4453	0.7967	0.5753	0.5902	2.4076



Cluster number/Algorithm : 4/Fuzzy C-mean
Iteration max. : 19
Average of the highest data traffic : 2.4412
Highest number of cluster members : 17
in months : 2

No.	Day	Rektorat	Fahutan	FMIPA	Fekon	Total
1	33	0.4262	0.7720	0.4697	1.0000	2.6679
2	35	0.5053	0.6652	0.3613	0.6528	2.1847
3	36	0.3931	0.8125	0.5078	0.2632	1.9766
4	37	0.4018	0.7949	0.5059	0.4293	2.1319
5	41	0.4310	0.8001	0.6145	0.6831	2.5286
6	42	0.5371	0.8739	0.5657	0.5641	2.5407
7	43	0.4459	0.7911	0.5324	0.5878	2.3572
8	47	0.4180	0.8140	0.6014	0.6816	2.5150
9	48	0.4072	0.9090	0.5361	0.4774	2.3296
10	49	0.3915	0.8273	0.5724	0.6154	2.4066
11	50	0.4167	0.7875	0.6895	0.4979	2.3916
12	51	0.3617	0.7700	0.5298	0.7608	2.4223
13	54	0.4249	0.8004	0.6672	0.6901	2.5826
14	55	0.5554	0.7653	0.5491	0.4991	2.3689
15	56	0.5460	0.8762	0.5908	0.3332	2.3462
16	57	0.5833	0.7647	0.9250	0.9108	3.1839
17	58	0.4497	0.9186	0.6768	0.5218	2.5669
Average		0.4526	0.8084	0.5821	0.5981	2.4412



Cluster number/Algorithm : 5/Fuzzy C-mean
Iteration max. : 14
Average of the highest data traffic : 2.4703
Highest number of cluster members : 16
in months : 2

No.	Day	Rektorat	Fahutan	FMIPA	Fekon	Total
1	33	0.4262	0.7720	0.4697	1.0000	2.6679
2	35	0.5053	0.6652	0.3613	0.6528	2.1847
3	37	0.4018	0.7949	0.5059	0.4293	2.1319
4	41	0.4310	0.8001	0.6145	0.6831	2.5286
5	42	0.5371	0.8739	0.5657	0.5641	2.5407
6	43	0.4459	0.7911	0.5324	0.5878	2.3572
7	47	0.4180	0.8140	0.6014	0.6816	2.5150
8	48	0.4072	0.9090	0.5361	0.4774	2.3296
9	49	0.3915	0.8273	0.5724	0.6154	2.4066
10	50	0.4167	0.7875	0.6895	0.4979	2.3916
11	51	0.3617	0.7700	0.5298	0.7608	2.4223
12	54	0.4249	0.8004	0.6672	0.6901	2.5826
13	55	0.5554	0.7653	0.5491	0.4991	2.3689
14	56	0.5460	0.8762	0.5908	0.3332	2.3462
15	57	0.5833	0.7647	0.9250	0.9108	3.1839
16	58	0.4497	0.9186	0.6768	0.5218	2.5669
Average		0.4564	0.8081	0.5867	0.6191	2.4703

Fig. 5. Plot of clustering network traffic results using fuzzy C-means

initial parameter values were implemented, (a). Number of clusters 3, 4, and 5; (b). Maximum iterations = 50; and (c). Stopping Criteria (ξ) = 10^{-5}. In this research, applied testing scheme is the amount of usage per day of data network traffic from the four parts (rectorate, forestry, science, and economic). Then, the

Table 2. Results comparison K-means and FCM methods

Cluster	Methods	Highest cluster members number	Average of network traffic usage	Highest of network traffic usage month
3	KCM	18	2.4076	2
	FCM	18	2.4076	2
4	KCM	14	2.5149	2
	FCM	17	2.4412	2
5	KCM	14	2.5149	2
	FCM	16	2.4703	2

highest, middle and lowest data values are obtained. The results of FCM are shown in Fig. 5.

In this study, the accuracy level in order to get the centroid value by using FCM method is quite careful than K-Means method, Table 2.

4 Conclusion

In this paper we presented a comparison of K-Means and FCM methods then compared the centroid accuracy by using various performance criteria. This research was used network traffic usage from four units; rectorate, forestry, science, and economic. In clustering, the examining of 1 parameter of centroid value is 3 centroid values. Our experiment showed that the FCM method is better results analysis in clustering than K-Means. Nevertheless, we have also concluded that FCM algorithm was slower than K-Means. As future work, an optimizing methods in order to get good accuracy between centroids is proposed.

References

1. Arroyo A, Herrero A et al (2016) Analysis of meteorological conditions in spain by means of clustering techniques. J Appl Logic. doi:10.1016/j.jal.2016.11.026.
2. Bai C, Dhavale D, Sarkis J (2016) Complex investment decisions using rough set and fuzzy c-means: an example of investment in green supply chains. Eur J Oper Res 248(2):507–521
3. Carvalho L, Barbon S Jr et al (2016) Unsupervised learning clustering and self-organized agents applied to help network management. Expert Syst Appl 54(C):29–47
4. Elaal A, Hefny H, Elwahab A (2010) Constructing fuzzy time series model based on fuzzy clustering for a forecasting. J Comput Sci 6(7):735–739
5. Huang CK, Hsu WH, Chen YL (2013) Conjecturable knowledge discovery: a fuzzy clustering approach. Fuzzy Sets Syst 221(2):1–23
6. Pandit YP, Badhe YP et al (2011) Classification of Indian power coals using k-means clustering and self organizing map neural network. Fuel 90(1):339–347

7. Stetco A, Zeng XJ, Keane J (2015) Fuzzy c-means++: fuzzy c-means with effective seeding initialization. Expert Syst Appl 40(21):7541–7548
8. Sucasas V, Radwan A et al (2016) A survey on clustering techniques for cooperative wireless networks. Ad Hoc Netw 47(C):53–81
9. Velmurugan T (2014) Performance based analysis between k-means and fuzzy c-means clustering algorithms for connection oriented telecommunication data. Appl Soft Comput 19(6):134–146