International Journal of Sustainability Management and Information Technologies
2019; 5(2): 39-44
http://www.sciencepublishinggroup.com/j/ijsmit
doi: 10.11648/j.ijsmit.20190502.13
ISSN: 2575-5102 (Print); ISSN: 2575-5110 (Online)

SciencePG
Science Publishing Group

Pixel Value Graphical Password Scheme: Compatibility of K-means Clustering Algorithm as Pixel Value Password Fault Tolerance Mechanism

Mohd Afizi Mohd Shukran, Mohd Sidek Fadhil Mohd Yunus, Muhammad Naim Abdullah, Mohd Nazri Ismail, Mohammad Adib Khairuddin, Kamaruzaman Maskat, Mohd Rizal Mohd Isa, Norshahriah Abdul Wahab, Mohd Fahmi Mohamad Amran

Department of Computer Science, National Defense University of Malaysia, Kuala Lumpur, Malaysia

Email address:

afizi@upnm.edu.my (M. A. M. Shukran), 3181079@alfateh.upnm.edu.my (M. S. F. M. Yunus), naim621@gmail.com (M. N. Abdullah), m.nazri@upnm.edu.my (M. N. Ismail), adib@upnm.edu.my (M. A. Khairuddin), kamaruzaman@upnm.edu.my (K. Maskat), rizal@upnm.edu.my (M. R. M. Isa), shahriah@upnm.edu.my (N. A. Wahab), fahmi@upnm.edu.my (M. F. M. Amran)

To cite this article:

Mohd Afizi Mohd Shukran, Mohd Sidek Fadhil Mohd Yunus, Muhammad Naim Abdullah, Mohd Nazri Ismail, Mohammad Adib Khairuddin, Kamaruzaman Maskat, Mohd Rizal Mohd Isa, Norshahriah Abdul Wahab, Mohd Fahmi Mohamad Amran. Pixel Value Graphical Password Scheme: Compatibility of K-means Clustering Algorithm as Pixel Value Password Fault Tolerance Mechanism. *International Journal of Sustainability Management and Information Technologies*. Vol. 5, No. 2, 2019, pp. 39-44. doi: 10.11648/j.ijsmit.20190502.13

Received: MM DD, 2019; **Accepted**: MM DD, 2019; **Published**: January 6, 2020

Abstract: In September 2018, the patent for pixel value graphical password scheme was granted in Malaysia. The graphical password scheme was designed to reduce the complexity of previously developed graphical password scheme where a user only requires to load their personal image as password instead of complex graphical challenge during authentication. As the guardian of digital access, Pixel Value Access Control was highly invincible from password pixel forgery attack where a little bit different pixel value derived from loaded image will deny the access. Only the original enrolled image from a registered user can be recognized by Pixel Value Access Control to authenticate the respective username. That fact makes the graphical password scheme is a trusted access control mechanism but, on the other hand, it makes users bound with the only original password pixel image file. Thus, Pixel Value Access Control need to be installed the pixel value fault tolerance mechanism where it could allow users to acquire their password pixel image file from various storage media. The clustering technique was suggested to solve this issue where it allows an altered pixel password being recognized as the same group of the original pixel password. There are number of clustering algorithms developed for various purposed and application of digital image clustering. K-Means algorithm is one the partition-based clustering algorithm that found to be the simplest and fastest clustering algorithm as suggested by many researchers. This paper is mainly to exhibit the selection of K-Means clustering algorithm became the crucial algorithm for Pixel Value Access Control password pixel fault tolerance algorithm. Background of this topic was briefly explained in introduction section, the implementation of K-Means algorithm as Pixel Value Access Control fault tolerance was elaborate in section 2 and followed by validation of the implementation in section 3. At the end of this paper, there is conclusion for this study and followed by list of references.

Keywords: PVAC, *PassPix*, K-means Algorithm, Graphical Password, Fault Tolerance, Euclidean Distance, Pixel Value, Image Query, Access Control

1. Introduction

The Pixel Value Access Control or shortly called as PVAC, [1] is an access control system that developed based on pixel value graphical password scheme method where it extract digital image file to validate a username. By using this access control system, users are required to upload their personally picked digital image file to the Pixel Value Access Control host server to being extracted its' pixel value. The pixel value is stored on the server alongside with the respective username as both data is used as the comparative query during authentication process. It was designed with such routine to simplify the authentication process while users dealing with graphical password. Previous graphical password scheme challenge users with series of image interaction which make hostile user partially hard to attempt but users also must deal with this tradeoff [2]. With Pixel Value Access Control, only users could simply use their credential image while others became inaccessible where the identity key is only hold by the actual users. The username authentication key is in the form as a digital image file that originates from users personal image storage provide by users during the enrollment and authentication process through the access control interface. As the digital image files hold a pixel value, Pixel Value Access Control extracts the pixel value as the password pixel, *PassPix*.

Figure 1. *The PVAC application runs on web-based.*

However, Pixel Value Access Control is intolerable to pixel alteration that caused users failed to authenticated and access denied. Since the concept of cloud storage has blend into society, internet users stored their data in cloud storage that allowed them data access flexibility [3] where it offer multiplatform application interfaces whether mobile-based apps or web-based application. Cloud storage media conditions and settings are uncertain where some of them are compressed and caused the pixel value alteration on *PassPix*. One of the common practices of cloud file repository is to transferred and stored files in WhatsApp [13, 3] messaging application where users could send, share and save their files in the app's cloud without draining their machine storage disk. Furthermore, this application is accessible from computer, desktop or laptop web-based, or through the smartphone apps, mobile-based apps. However, WhatsApp application does compress every image-based multimedia files, pictures and videos, from the moment it being upload to the application to ensuring the message delivery speed and server side storage. In order to restore the boundless to a single image file for users, the *PassPix* fault tolerance mechanism is a must for Pixel Value Access Control. Adapting a clustering technique into Pixel Value Access Control is a possible way as fault tolerance mechanism where the altered *PassPix* is recognized as the same category.

2. The Clustering Techniques

Image clustering or segmentation is referring as an automated process to categorizing a collection of images into few categories based on the features or content of an image. There are 3 type of feature recognition for clustering techniques which are:

i. Edge based recognition.
ii. Region based recognition.
iii. Pixel based recognition.

The Pixel-based clustering technique is simple structured technique consuming least computational resources and produced an acceptant segmentation output [5]. It start with pixel extraction of an image based on color where the color scheme is define by user, for example RGB color scheme. Then the extracted pixel object is placed into their categories based on clustering algorithm. The process of placing the extracted object called as membership assigning. It employs 2 type of membership assigning methods [6] which are hard clustering and soft clustering. Hard clustering method assigned an extracted object to one cluster only which employ the rule one object must be belong to one cluster only. While the soft clustering assigned an extracted object to more than one cluster which employ the rule one object could be belong to more than one cluster as briefly illustrated in figure 2.

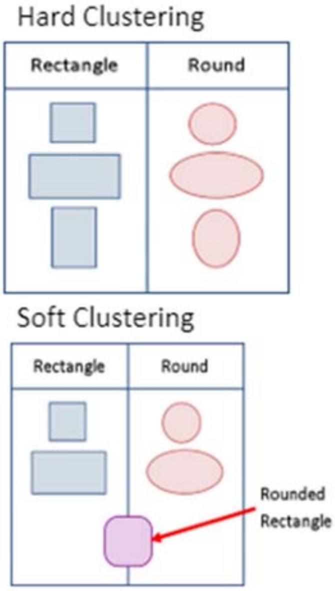
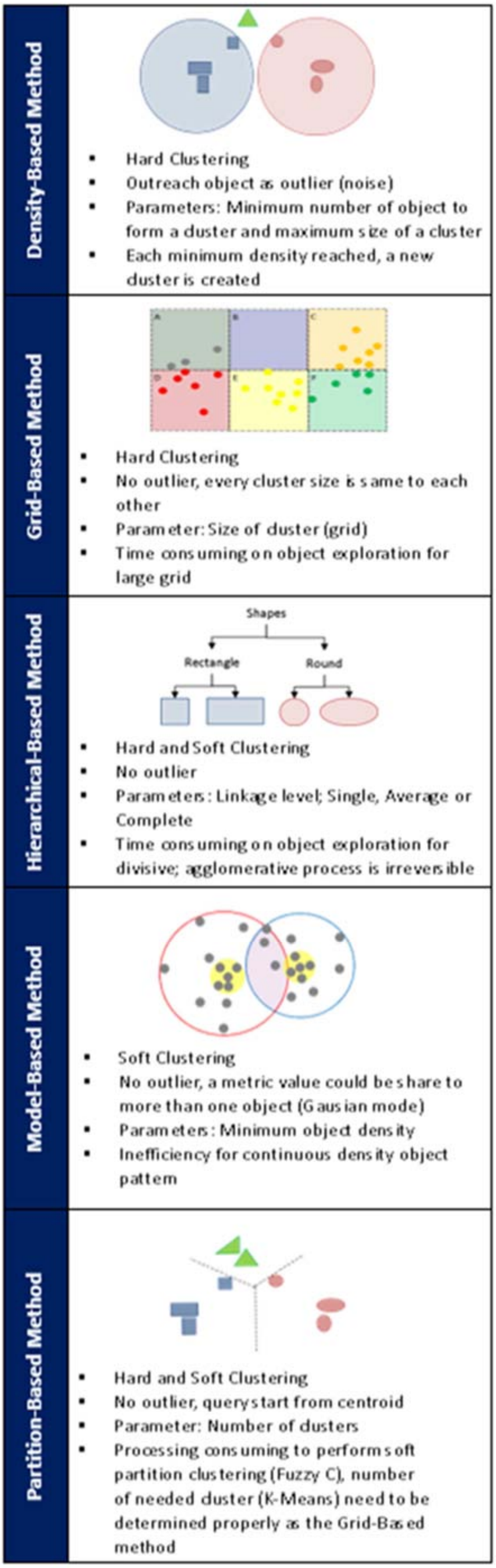


Figure 2. *The cluster membership assigning method.*

When each object was successfully resided on its cluster respectively, the query for each object is compare on its clusters membership. In a cluster the query method employ is Euclidean distance [7] based on the clustering pattern. The clustering pattern can be divided into 5 methods [8, 9] which are:

i. Density-Based method.
ii. Grid-Based method.
iii. Hierarchical-Based method.
iv. Model-Based method.
v. Partition-based method.

The performances of the query process and resource consumption are based on the clustering pattern. The details characteristics and the different for each clustering method is briefly illustrate in figure 3.



Density-Based Method

- Hard Clustering
- Outreach object as outlier (noise)
- Parameters: Minimum number of object to form a duster and maximum size of a cluster
- Each minimum density reached, a new duster is created

Grid-Based Method

- Hard Clustering
- No outlier, every cluster size is same to each other
- Parameter: Size of duster (grid)
- Time consuming on object exploration for large grid

Hierarchical-Based Method

- Hard and Soft Clustering
- No outlier
- Parameters: Linkage level; Single, Average or Complete
- Time consuming on object exploration for divisive; agglomerative process is irreversible

Model-Based Method

- Soft Clustering
- No outlier, a metric value could be share to more than one object (Gausian mode)
- Parameters: Minimum object density
- Inefficiency for continuous density object pattern

Partition-Based Method

- Hard and Soft Clustering
- No outlier, query start from centroid
- Parameter: Number of dusters
- Processing consuming to perform soft partition clustering (Fuzzy C), number of needed duster (K-Means) need to be determined properly as the Grid-Based method

Figure 3. *The 5 cluster pattern characteristics and different.*

K-Means algorithm [10] as an algorithm that categorized in partition-based method, is nominated as the fault tolerance mechanism for Pixel Value Access Control due to its simplicity that would execute clustering faster than other algorithms [11]. This characteristic of K-Means performance is highly appropriate due to Pixel Value Access Control need to ready to handle multiple request of *PassPix* query in a single time since it is function as the frontline of access. For such requirement, Pixel Value Access Control require a clustering algorithm for that consumed least resources & time for the *PassPix* query where K-Means performed object exploration faster and require minimal processing resources for clustering as stated in figure 3.

By the concept, K-Means performance is an appropriate match for Pixel Value Access Control. However, only few parameters that exceptionally useful for Pixel Value Access Control where:

i. The tolerance value need to be fixed to avoid the *PassPix* value mismatch.
ii. The tolerable range is crucial where if the range is too wide the *PassPix* became less exclusive.
iii. RGB – based data extraction and clustering capabilities.

Thus, there are 2 values that could adopted for Pixel Value Access Control which are:

i. Cluster Vector; the cluster ID of an object.
ii. The Object Distance (Euclidean Distance); the similarity range of a query object with the clustered objects.

However, the cluster vector is unhinge for each object added to the database which caused the vector value mismatch and Pixel Value Access Control will denied the access. There are number of other clustering algorithm that could fixed this error but, the Pixel Value Access Control authentication process performance need to be sacrifice that consume a lot more resources and time. The only remaining value that suitable for Pixel Value Access Control is Euclidean distance which needs to be set just nice to avoid the fake *PassPix* toleration. The reliable value for Pixel Value Access Control is the Euclidean distance and the distance parameter value is possible implementation without overhauling the clustering algorithm. For this issue, it will be discovered in our further experiment on K-Means for Pixel Value Access Control and just for this topic, we perform the validation based on algorithm performance.

3. Validation of K-Means Compatibility for Pixel Value Access Control

In this study, the compatibility validation for K-Means and Pixel Value Access Control is performed as a comparative experiment with another clustering algorithm. STING [12] is categorized as one of the grid-based cluster patterns is used as the sparing algorithm with K-Means. Besides its cluster pattern, the characteristic of clustering rules where both K-Means and STING are hard clustering, leave no outlier vector data and number of cluster parameter setting. The only feature that differentiates STING from K-Means is the clustering pattern that will affect the object classification as well as the query process. For this validation study, the STING algorithm is load into the clustering application and the data set used is 2,000 image acquired from CALTEX 101 object categories [4] which customly categorized into 10 categories and 20 categories. The lab configuration is briefly illustrated in figure 4.

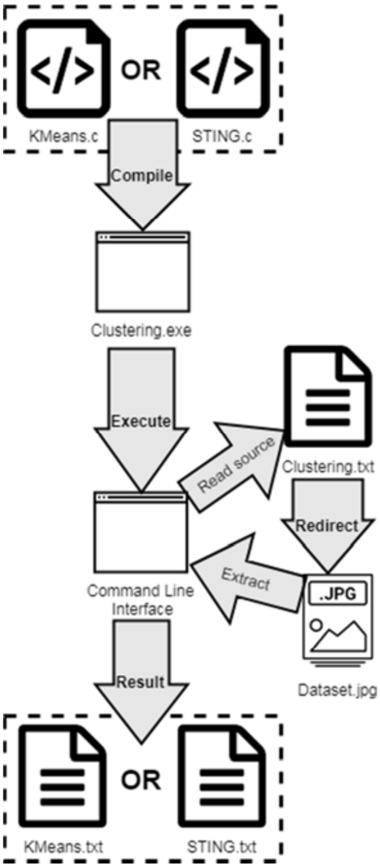


Figure 4. *Lab environment for the K-Means vs STING.*

The comparative validation between K-Means and Sting is using a clustering application that develops in C++ language and can be equipped with the clustering algorithm script either K-Means or STING during compilation. Each clustering algorithm is compile separately and execute separately. The interface of the application is command line interface which is appearing when the clustering application is executing. During the clustering process, the list of data set is acquire from clustering dataset text file that redirect the application to the image directory that contain 2,000 undivided image file which all 2,000 images placed into a single directory. There are 2 types of image dataset where it is custom picked based on image categorization by the dataset provider, CALTEX 101. The first dataset is consist of 20 category images where each category is consist of 100 image files. The second dataset is made up from 40 category images and each category consists of 50 images. That's make both categories is consist of 2,000 images. All of the images is extracted by the clustering application to acquire the pixel value in the form of RGB strength. The clustering application is clustering the RGB value and the complete result is produced on a text file. When the process completed, the time taken for the clustering process is recorded as the comparative value for all of the testing. There are 4 testing in total performed in this study. The testing process is briefly illustrated in figure 5 and example of the result produced shows in figure 6.

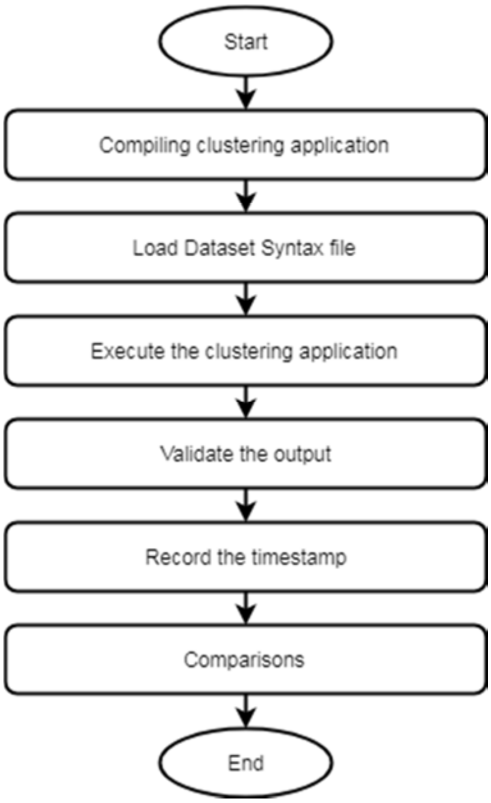
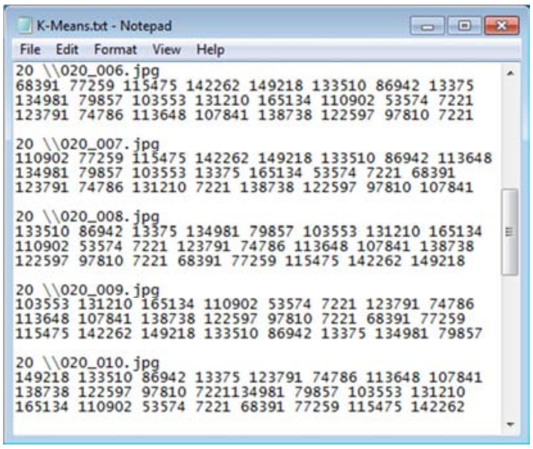


Figure 5. *Clustering validation process for K-Means vs STING.*



```
20 \\020_006.jpg
68391 77259 115475 142262 149218 133510 86942 13375
134981 79857 103553 131210 165134 110902 53574 7221
123791 74786 113648 107841 138738 122597 97810 7221

20 \\020_007.jpg
110902 77259 115475 142262 149218 133510 86942 113648
134981 79857 103553 13375 165134 53574 7221 68391
123791 74786 131210 7221 138738 122597 97810 107841

20 \\020_008.jpg
133510 86942 13375 134981 79857 103553 131210 165134
110902 53574 7221 123791 74786 113648 107841 138738
122597 97810 7221 68391 77259 115475 142262 149218

20 \\020_009.jpg
103553 131210 165134 110902 53574 7221 123791 74786
113648 107841 138738 122597 97810 7221 68391 77259
115475 142262 149218 133510 86942 13375 134981 79857

20 \\020_010.jpg
149218 133510 86942 13375 123791 74786 113648 107841
138738 122597 97810 7221134981 79857 103553 131210
165134 110902 53574 7221 68391 77259 115475 142262
```

Figure 6. *Screenshot of the result for the K-Means 40 categories.*

As shown in figure 6, the first line on each object indicates the cluster ID and followed by the filename. The rest of the data is the pixel value or RGB value for the respective image. CALTEX 101 dataset assigned the filename as:

[Category Number]_[File Number].jpg.

This indicated that both algorithm performed the clustering accurately where all images was categorized accordingly. However, time taken for all testing is different. The result of all 4 testing is summarized in table 1.

Table 1. *Clustering time taken for K-Means vs STING.*

Number of Clusters	K-Means	STING
20	13 Seconds	19 Seconds
40	22 Seconds	31 Seconds

The result shows that, K-Means perform faster than STING as mentioned in the concept of adapting K-Means for Pixel Value Access Control in the previous section. The time taken is significantly increase when the number of cluster increase for both algorithm. However, K-Means recorded less increment compared to STING. There are only 9 seconds increments for K-Means compared to 12 seconds for STING. For additional categories, STING will consume lot more time to clustering objects as shown in the forecast graph, figure 7.

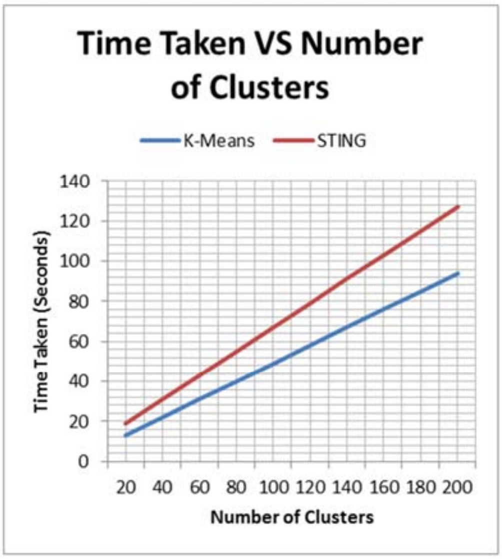


Figure 7. *Time taken against number of clusters for K-Means and STING.*

Even though both algorithms consume more time as the number of cluster increased, STING recorded steeper line pattern compared to K-Means. This to validate that, K-Means is a more reliable to process large number clustering compares to STING. For that K-Means characteristic is useful for Pixel Value Access Control application when dealing with large number of clustering request from the users. From that finding, K-Means is compatible to be adapted into Pixel Value Access Control that highly possibility must deals with a lot of authentication request in same time.

4. Conclusion

Since Pixel Value Access Control is dealing with pixel value extraction to authenticate a user, Pixel-Based clustering technique is adapted for Pixel Value Access Control fault tolerance mechanism. This clustering technique is using the same pixel value extraction method as Pixel Value Access Control extraction method where RGB was chosen as the pixel value color scheme. There are no modification or additional pixel value extraction process need to be added into Pixel Value Access Control that would deficiency of current performance of Pixel Value Access Control. This will allow the Pixel Value Access Control to accept and authenticate a password pixel that has been altered unintentionally due to many reasons. It is an obvious finding that K-Means clustering algorithm is the most compatible clustering algorithm to be adapting into Pixel Value Access Control as the fault tolerance mechanism. From the testing that exhibit in this study, K-Means outperformed STING in performing the clustering process for 2,000 images. However, as shown in forecast graph figure 7, both algorithm time consuming to perform clustering process increase significantly when the number of cluster is increasing. It is advisable while designing any image clustering related program to reduce the number of cluster and make it as minimal as possible to avoid high consumption of time and resources. Time consumption and processing resource consumption are relative where longer time consumption also means high processing resource consumption.

This study is only a part of designing and adapting a clustering technique as Pixel Value Access Control pixel fault tolerance. While the clustering process is a conceptual method to speed up the query process wile calculating pixel value differences, the effect of K-Means clustering algorithm on the pixel value distance query yet to be discover. Our tremendous effort on developing and enhancing Pixel Value Access Control is far from the end. There are more room for improvement for Pixel Value Access Control in the future. We hope that Pixel Value Access Control soon will become a major hit for graphical access control implementation.

References

[1] M. A. M. Shukran & M. S. F. M. Yunus, “Method and System For Authenticating User Using Graphical Password For Access Control,” Malaysian Patent MY-167835-A, 2018.

[2] M. S. F. M. Yunus, "Dynamic Analysis of Pixel Value Graphical Password Scheme," Master Thesis, National Defense University of Malaysia, Kuala Lumpur, 2014.

[3] M. A. M. Shukran, M. S. F. M. Yunus, M. N. Abdullah, M. N. Ismail & M. R. M. Isa, "Pixel Value Graphical Password: A PassPix Clustering Technique For Password Fault Tolerance," International Journal of Recent Technology and Engineering (IJRTE) ISSN: 2277-3878, Volume-8 Issue-3, September 2019, pp. 2973-2976.

[4] L. Fei-Fei, R. Fergus & P. Perona, "Learning generative visual models from few training examples: an incremental Bayesian approach tested on 101 object categories", IEEE. CVPR 2004, Workshop on Generative-Model Based Vision. 2004.

[5] M. Panda, A. E. Hassanien, & A. Abraham, "Hybrid Data mining approach for image segmentation based Classification," Biometrics: Concepts, Methodologies, Tools, and Applications, 2017, pp. 1543-1561.

[6] B. A. Pimentel, & R. M. Souza, "Multivariate Fuzzy C-Means algorithms with weighting," Neurocomputing, 174, 2016, pp. 946-965.

[7] E. Schubert, J. Sander, M. Ester, H. P. Kriegel & X. Xu, "DBSCAN Revisited, Revisited: Why and How You Should (Still) Use DBSCAN," ACM Trans. Database Syst. 42 (3): 19:1–19:21. doi: 10.1145/3068335, 2017, pp. 0362-5915.

[8] A. Fahad, N. AlShatri, Z. Tari, A. Alamri, I. Khalil, A. Y. Zomaya & A. Bouras, "A Survey of Clustering Algorithms for Big Data: Taxonomy and Empirical Analysis," IEEE transactions on emerging topics in computing, 2 (3), 2014, pp. 267-279.

[9] P. Sharma, & J. Suji, "A review on image segmentation with its clustering techniques," International Journal of Signal Processing, Image Processing and Pattern Recognition, 9 (5), 2016, pp. 209-218.

[10] J. Macqueen, "Some methods for classification and analysis of multivariate observations," Proceedings of the fifth Berkeley symposium on mathematical statistics and probability, 1, 1967, pp. 281-297.

[11] K. Rajalakshmi, D. S. Dhenakaran, & N. Roobin, "Comparative Analysis of K-Means Algorithm in Disease Prediction," International Journal of Science, Engineering and Technology Research (IJSETR), 4 (7), 2015, pp. 1-3.

[12] W. Wang, Y. Jiong & M. Richard, "STING: A statistical information grid approach to spatial data mining," VLDB. Vol. 97, 1997.

[13] WhatsApp Inc., WhatsApp Features, Available: https://www.whatsapp.com/features/, 2019.