

Perception of students on usage of mobile data by K-mean clustering algorithm



Arshad Khan¹, Samad Baseer², Sarfaraz Javed^{3,*}

¹Department of Computer Science, Qurtuba University of Science and Technology, Peshawar, KPK, Pakistan

²Department of Computer System Engineering, University of Engineering and Technology, Peshawar, KPK, Pakistan

³Department of Commerce, Aligarh Muslim University, Aligarh, India

ARTICLE INFO

Article history:

Received 19 October 2016

Received in revised form

10 December 2016

Accepted 13 December 2016

Keywords:

M-learning

3G/4G mobile technology

K-mean clustering algorithm

Higher education

ABSTRACT

This study is conducted in order to analyze the perception of students on use of mobile learning 3G/4G in higher education. Data has been collected through the survey data, this study comprises research questions based on the previous literature review and formulated hypotheses to check the tentative statements. Furthermore, collected data has been analysis using K-mean algorithm. Additionally, this study explore methodological gap using K-Mean clustering algorithm. The sample study consisted of 200 students in the department of computer science at Gomal and Qurtuba universities in Dere Ismail Khan.

© 2017 The Authors. Published by IASE. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

1. Introduction

Wide spreading of the fame of the Mobile Data (3G/4G) in youth is expanding at Higher Education level (Brown, 2003). For this kind of learning, understudies and educator do not meet in personal and there is no particular time when the students meet for learning. For this strategy instructors or teachers speak with their students by means of telephones, PC, web and so forth such learning is typically take place with the utilization of electronic and print media like recordings, messages, daily papers and so on (Wendeson et al., 2010). This paper takes a gander at the significance of learning and training can be enhanced by the utilization of 3G/4G particularly in advanced education. It will also solve its statement of the problem by looking and different clusters which have been clustered using K-Mean clustering algorithm (Caladine, 2008).

1.1. Statement of problem

There is a rise in popularity of mobile data 3G/4G learning in most the institution of higher learning. Most universities are using online services that is because not only for registration but also in all their units. It uses in some full courses and that are done

online from the first to the final year (Sharples et al., 2005). Most institution is starting to adapt this technology and some even go to extents of spending a lot in the project that ensure proper covering of mobile data (Sharples et al., 2005). To ensure the success of this mobile learning, students have to accept this technology since they are the only targets for this technology (Carlson, 2005). This paper seeks to know how student perceives the application of ICT, especially the 4G/3G mobile data technology. Understanding how the students perceive this will go a long way in understanding what needs to be done; more investment on it or stop the investment and begin sensitizing the student on it. It will look at some of the factors that affect 3G/4G mobile learning technologies in the institution of higher learning. When an institution is able to know what the student think of mobile data 3G/4G, they will be able to know where to invest and not invest especially when it comes to social media. More research in this can act as a very good asset to institutions (Attewell and Savill-Smith, 2004). There are many factors that affect this; they include availability, ease of use, affordability, speed and accessibility. However, the study consist two variables, which is availability and ease of use. This will seek to answer of questions like, is mobile data available for use by the most student? Well even if it is available, what do student think of its usability? Once these questions are answered the researcher will have collected the expected data.

* Corresponding Author.

Email Address: sarfaraz.rs@amu.ac.in (S. Javed)

<https://doi.org/10.21833/ijaas.2017.02.003>

2313-626X/© 2017 The Authors. Published by IASE.

This is an open access article under the CC BY-NC-ND license

(<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

1.2. Research questions

Most the research question and hypothesis being tested in what the student think of the 3G/4G influence on educations. There are three research questions that the study will seek to answer:

- Q1: what is the Perception of undergraduates and graduates on the implementation of mobile learning technologies for learning in selected institutions in Dera Ismail Khan.
- Q2: what is the Perception of undergraduates and graduates on the availability and the ease of use of mobile learning technologies?
- Q3: what do student think of mobile data 3G/4G, is it always available for use when they need it?
- Q4: what do students think of its use, is it easy to use and is there enough training for the same?
- Q5: what do students think, are there other opportunities where mobile data can invest into in terms of education?

1.3. Hypothesis

H1. Graduates think and perceive mobile data 3G/4G as an advantage to most institution when used in the learning process.

H2. Most students believe that the implementation of mobile data 3G/4G will improve the condition of students in the university

H3. Graduates and undergraduates think that mobile data 4G/3G has a long way to go before all people are able to accommodate it in the learning process

H3. There is not feature for mobile data especially in the field of graduate and undergraduate educations

2. Methodology

2.1. Research design

The paper has used the descriptive method of design for its research in acquiring personal data. This means that the study describes the way things are and provides detailed answers to questions and comments. It uses quantitative investigation methods to collect data. Data will be collected through observations, surveys, questionnaires and interviews. To develop quantitative or qualitative data most of the researchers use a way which is known as the questionnaire (Sekaran, 1999). The stakeholder's answers are recorded or written and can be used to represent the stakeholder's preferences. Questionnaires will be used as they are time capable, economical and in most cases allow the researcher sort out questions and find the huge number and figures without having to talk to every participant (Nazeer and Sebastian, 2009).

To make sure the validation of questionnaire, the researcher used adopted questionnaire consisting of closed-ended 31 questions. These questions were

used in the previous research to find out about the mobile devices and on internet usability and the cost internet when being used for learning (Collis et al., 2013).

2.2. Data collection methods

Data were collected from both secondary and primary sources. Secondary sources included books, journals and other written sources. Primary sources for this research paper included interviews, managing questionnaire and observing individuals. The methods of primary data collection that will be used are a questionnaire, interviews and field observations (Nazeer and Sebastian, 2009). A semi-structured questionnaire will be developed and administered to respondents in 20 to 30 minutes depending on the patterns, speed and comprehension and clarity of response. This is done through face to face interview with respondents. The questionnaire contains different sections i.e. background information and specific questions relevant to the study.

Pretesting were carried out so as to provide an opportunity to see what questions work well, which questions can be eliminated and which need to be added, see whether questionnaire is too long and boring to the respondents and whether the questions are simple and understandable. Pretesting has done to my follower students to take note of some problematic words and flow of questions then after revising my questionnaire. Pre-test it again within the surrounding community who have similar characteristic on the level of education as my pre-assumed population.

2.3. Sample size determination

A cross-sectional survey which is use structured questionnaire. According to Kombo and Tromp (2006) survey design is appropriate for collecting analyzing, comparing, and interpreting data .There presentative sample size was determined by the Krejcie and Morgan (1970) formula commonly used to calculate a simple size from a given finite population (P) such that a sample size is within plus or minus 0.1of the population within a 95% level of confidence.

The sample study consisted of 200 students in the department of computer science at Gomal and Qurtuba universities in Dere Ismail Khan.

2.4. Likert scale

Likert scale used to evaluate the student perception on mobile data 3G/4G in institutions of higher learning. Likert scale is one of the major universal techniques used for conducting research (Switzer and Csapo, 2005; Trifonova et al., 2006). It is a combination of elements which focus on a particular problem or idea. It employs coding for some of the questions, for example, when a

statement like '3G/4G mobile technology is available for learning all time' the Likert scale will give codes of 1= strongly agree 2= Agree 3 = undecided 4 = Disagree and 5 equal to Strongly Disagree (Fetaji et al., 2011).

2.5. Data clustering

Data clustering is an unsupervised and statistical data analysis technique which is used to classify or cluster the different data sources into one homogenous group (Costa et al., 2008). The role of data clustering in this study is used to operate the large data sets as it is used to uncover relationships and patterns to help make a decision quickly and efficiently. Nonetheless, it can be used to cluster small and average data sets to create confidence in the study.

2.6. K-means clustering algorithm

K-Means is a simple unsupervised learning algorithm that is used in clustering. K- Mean partition and 'n' is the number of observation into k

clusters with each observation belonging to the cluster with the nearest mean. The k-mean algorithm aims at minimizing the objective function. For our data, the process of clustering data using K-algorithm is as follows:

1. Dividing the dataset into K-clusters and the assigning the clusters at random. As a result, it forms almost the same number of the data point on them. The random cluster makes sure that each and every entry point has the same probability of being clustered in any clusters. This will avoid biases.

2. The distance from one data point to each cluster is then assessed for every point.

3. The data point not near to any cluster is then moved to a particular cluster that is near; the data point near to a cluster is left alone.

4. The final clusters that result in terms of initial cluster and intra-cluster distances and cohesion can be extremely affected by the choice of initial partition (Sangeetha et al., 2015).

The flow chart for k-mean algorithm is shown in Fig. 1.

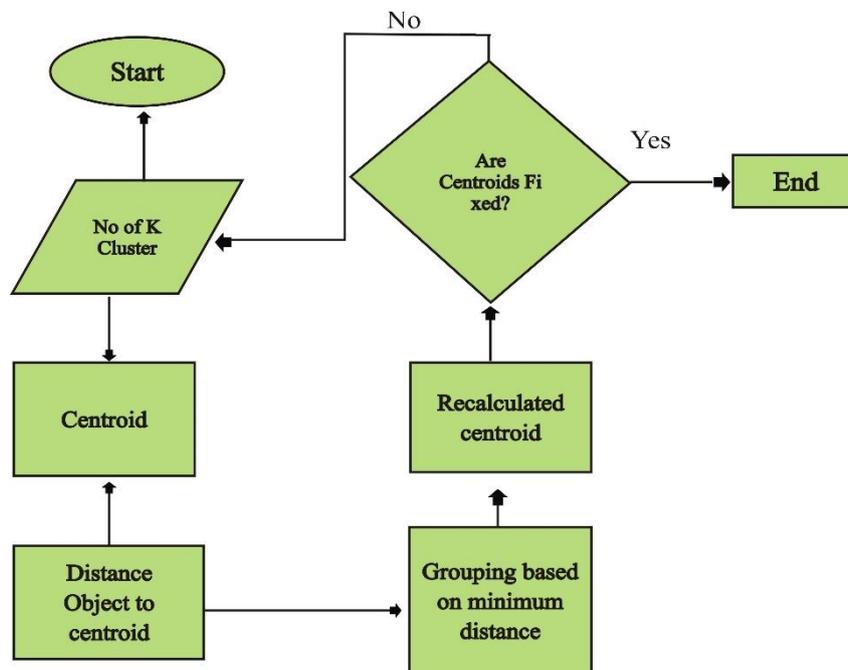


Fig. 1: Flow chart of K-mean algorithm

3. Data analysis

The data analyzed using both qualitative and quantitative techniques. SPSS version 20 were used for analysis. This data include age and gender of the respondents, educational level etc. Descriptive statistics also used, such as a measure of central tendency e.g. percentages and frequencies. After that k-mean clustering analysis were used for dividing the data into clusters. Inferential statistics were used to measure the relationship between variables i.e. ANOVA. This is the most appropriate models for the independent variables are qualitative (dummy variables) and quantitative. To present the data,

graphs, pie charts, ANOVA table and table of frequencies will be used. ANOVA is the Analysis of Variance and the original thinking is that is used to try and partition the variance in response to all variances, factors and errors (Costa et al., 2008) (Table 1).

4. Findings

Descriptive Statistics Table introduces the frequencies between respondents of gender orientation, age, institute, program, and program year.

The above Table 1 is showing the total number of respondents included in this study with minimum and maximum values. To get to the clusters, we use mean values. However, to be more specific and get deeper into the clusters, we used standard deviation

feature to get the variance between the different points. As Table 1 shows, we get the mean and then from the mean, we get the standard deviation to show the variance in the means Analysis was done through ANOVA (Table 2).

Table 1: Descriptive statistics

	N	Minimum	Maximum	Mean	Std. Deviation	Variance
Age	181	1.00	3.00	1.6077	.69582	.484
Gender	181	1.00	2.00	1.3039	.46120	.213
Institute	181	1.00	2.00	1.5304	.50046	.250
Program	181	1.00	3.00	1.5746	.71586	.512
Year of program	181	1.00	4.00	2.1934	1.01169	1.024
Valid N (list wise)	181					

Table 2: Anova table

	Cluster		F	Sig.
	Mean Square	df		
Availability	20.613	2	83.595	.000
Ease of use	18.554	2	74.377	.000

4.1. ANOVA analysis

The ANOVA table was to indicate the variables that contributed most to the cluster solution. It is clear that those that had a large F value that is Easy of use, availability provided the same values between the clusters. It is also clear that all the parameters are significant in explaining the model since their p values are less than 0.05 level of significance.

4.2. Case processing summary

Both undergraduate and graduate students strongly agree that the mobile technology is accessible to every student. Data showed that most of the students agree on the availability and ease of use of the mobile technology in their universities. A minority of the students both graduate and undergraduate strongly disagree on the availability and ease of use of the mobile technology. The majority of the students both graduate and undergraduate agree that the mobility technology is cost effective for the student that is 53 students and 50 students agreed on the same that can be interpreted as 29.3% and 27.6% respectively. 34 (18.8%) students were undecided 41(22.7%) disagreed and 3 (1.7%) disagreed. 72 students (39.7%) both graduate and undergraduate agreed that 3G/4G mobile learning technologies give rapid access to course related material because of its high speed, 63 students (34.8%) strongly agree, 12 (6.6%) were undecided, 16 (8.83%) disagreed and 18 (9.9%) strongly disagreed on that (Costa et al., 2008).

4.3. K-mean implementation

The final cluster centers were computed as the mean for each variable within each cluster and they reflect the characteristics of each case for each cluster (Fig. 2 and Table 3).

Students in Cluster 1, strongly believe that 3G/4G mobile technology is available and accessible but not

affordable for them though they tend to believe that the 3G/4G mobile technology provide quick way to access online learning materials and it helps to accomplish their studies inconvenient time and mobile technology enhance learning better than adopted technologies.

Table 3: Final cluster centers

	Cluster		
	1	2	3
Availability	2.74	1.79	2.90
Ease of use	2.52	1.91	3.84

Students in Cluster 2 agree that the 3G/4G mobile technology is available and accessible, but can be affordable for some of them. They also agree that mobile technology should be implemented in the learning institutions since it helps the student accomplish their studies at a convenient time.

Final Cluster Centers

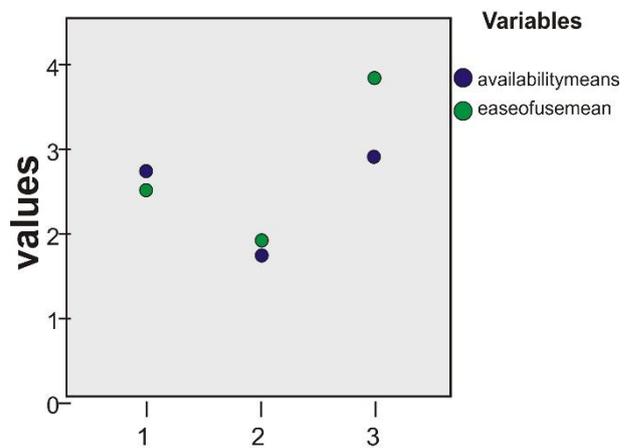


Fig. 2: Final cluster centers

Students in cluster three less agree on the availability and Ease of use of the 3G/4G mobile technology though they think it can be affordable to the students (Table 4).

Table 4: Iteration history

Iteration	Change in Cluster Centers		
	1	2	3
1	1.577	1.296	1.352
2	.111	.090	.000
3	.000	.000	.000

The iteration table shows difference between clusters mean values. Initially the mean values are same and high between the clusters. After each iteration the mean values decreasing in each cluster (Table 5).

Table 5: Distances between final cluster centers

Cluster	1	2	3
1		1.125	1.335
2	1.125		2.228
3	1.335	2.228	

5. Discussion and conclusion

For the k-mean cluster analysis procedure the students were grouped in three clusters, it is difficult to know the best number of clusters to choose when performing the K-mean analysis but when you rerun analysis with different number of clusters and examine the solutions you come up with the most appropriate grouping. For this case the three clusters gave the best results and explained the groups in a better way. From the analysis we can conclude that most of the students agreed on the fact the 3G/4G mobile technology can be available and accessible to most of the students. The K-mean analysis is important in making sure that the squared Euclidean distance between the clusters is minimized. In our case the students both graduate and undergraduate are grouped in three clusters and each cluster has students with close characteristics. The results showed that both graduate and undergraduate in the selected universities said that the availability and the ease of use of 3G/4G mobile technology were indirectly proportional. Technology is readily available; there are not enough people to teach people on its ease of use. However, a majority of the student said that it should be implemented in all universities as it has a lot of advantages. The study came up to a conclusion that student perceives 3G/4G in a positive way and think that it should be expanded even more.

References

- Attewell J and Savill-Smith C (2004). Learning with mobile devices: Research and development: A book of papers. Learning Skills and Development Agency, London, UK.
- Brown TH (2003). The role of m-learning in the future of e-learning in Africa. In 21st ICDE World Conference, Hong Kong. Available online at: <http://www.tml.tkk.fi/Opinnot/T-110.556/2004/Materiaali/brown03.pdf>
- Caladine R (2008). Enhancing e-learning with media-rich content and interactions. Information Science Publishing, New York, USA.
- Carlson S (2005). The net generation goes to college. The chronicle of higher education, 52(7): A34. Available online at: <http://www.chronicle.com/article/The-Net-Generation-Goes-to/12307>
- Collis J and Hussey R (2013). Business research: A practical guide for undergraduate and postgraduate students. Palgrave Macmillan, Basingstoke, UK.
- Costa JC, Ojala T, and Korhonen J (2008). Mobile lecture interaction: Making technology and learning click. In IADIS International Conference Mobile Learning, Algarve, Portugal: 119-124.
- Fetaji B, Ebibi M, and Fetaji M (2011). Assessing effectiveness in mobile learning by devising MLUAT (mobile learning usability attribute testing) methodology. International Journal of Computers and Communications, 5(3): 178-187.
- Kombo DK and Tromp DL (2006). Proposal and thesis writing: An introduction. Nairobi: Paulines Publications Africa, Nairobi Kenya.
- Krejcie RV and Morgan DW (1970). Determining sample size for research activities. Educational and Psychological Measurement, 30: 607-610.
- Nazeer KAA and Sebastian MP (2009). Improving the accuracy and efficiency of the k-means clustering algorithm. In the Proceedings of the World Congress on Engineering. Available online at: http://www.iaeng.org/publication/WCE2009/WCE2009_pp308-312.pdf
- Sangeetha K, Mohanapriya D, Gowathami D, and Mujeeb M (2015). Web usage data clustering using improved k-means algorithm. International Journal of Advanced Research in Computer Science and Software Engineering, 5(8): 242-245.
- Sekaran U (1999). Research methods for business: A skill-building approach. 3rd Edition, Wiley, New Jersey, USA.
- Sharples M, Taylor J, and Vavoula G (2005). Towards a theory of mobile learning. In the Proceedings of MLearn, 1(1): 1-9. <http://www.compassproject.net/sadhana/teaching/readings/sharplesmobile.pdf>
- Switzer S and Csapo N (2005). Survey of student usage of digital technology: Teaching implications. Issues in Information Systems, VI(1): 127-133.
- Trifonova A, Georgieva E, and Ronchetti M (2006). Determining students' readiness for mobile learning. In Proceedings of the 5th WSEAS International Conference on E-ACTIVITIES. Venice, Italy: 84-89.
- Wendeson S, Ahmad WFBW, and Haron NSB (2010). University students awareness on M-learning. World Academy of Science, Engineering and Technology, International Journal of Social, Behavioral, Educational, Economic, Business and Industrial Engineering, 4(2): 133-137.