A NEW SCHEME OF SCHOOL CLUSTERING FOR BETTER MANAGEMENT: A STUDY ON GRADE V SCHOLARSHIP EXAMINATION RESULTS IN GAMPAHA DISTRICT

RMATK Rathnayake[#] and RAB Abeygunawardana
Department of Statistics, Faculty of Science, University of Colombo, Sri Lanka

[#]Anu1991rathnayake@gmail.com

Abstract—*Educational management is some kind of complex human enterprise in order to achieve and to accomplish the expected goals by bringing different resources together. School management, as a fundamental of educational management, consists number of principles relating to the technique of classroom procedure and the practice of successful teachers. Hence most of the developed countries tend to use concept called "Clustering schools". In Sri Lanka, government schools are classified into four types namely type 1AB, type 1C, type 2 and type 3 based on streams of advanced classes and not based on the academic performance of students. This study is based on the students' performance using Grade-v scholarship examination results in year 2013 and 2014. This proposed classification method can be used to find out best schools, lower performance schools as well as moderate schools in the particular area. By identifying school performance, it may be helpful to authorized people in education in order to introduce new methods to improve the academic performance of schools. Therefore they can manage and update proper educational system for upcoming years.*

Schools in Gampaha district were selected as a sample which contains four zones namely Minuwangoda, Kelaniya, Gampaha and Negombo. Cluster analysis was carried out in each zone separately under criteria namely boys/mixed schools, girls/mixed schools and mixed schools. The application of hierarchical method (wards method, single linkage, average linkage and complete linkage) was used for clustering and three clusters were defined in each method. According to the result, same set of schools in Gampaha, Kelaniya, Negombo and Minuwangoda zone were identified as best schools and low performance schools in both years. However, some of the schools that belong to the Moderate group were moved to the Best and Low performance groups in two years.

Keywords—*Clustering, schools, hierarchical, management*

I. INTRODUCTION

Education is essential for every living being and importance of education is undeniable for every single person. Because of that it generates an advance community and even imagines a life without it. Even though education has influenced on life, still there are some places in the world where most of people are completely uneducated. According to United Nation's Children Fund (UNICEF) point of view, there are number of positive effects when providing good education for children such as decreasing poverty, decreasing child mortality rate, encouraging environmental concern and so on.

Sri Lankan education system provides free education for children who are in state schools in Sri Lanka. As a developing country this is a great victory for all students and this is a great opportunity that they are able to have school education as well as higher education without any cost.

This study is based on Grade-v scholarship examination which is highly competitive examination in Sri Lankan education system conducted by the Department of Examinations of the Ministry of Education. At the end of the primary education, (ages 9-10) students in Grade-V are sat to this scholarship examination which is conducted Ministry of Education, Department of Examinations in each year. It is optional for those students and they can transfer to national schools in urban cities based on their results.

This study is aimed to achieve objectives follows.

- To identify performance of Grade-v scholarship results of schools in Gampaha district
- To build a classification scheme for schools.

Proposed classification method can be used to find out not only the best schools in the particular area, but also lower

performance schools in that area. If all schools have same in performance of education, there will not be any competition among parents when entering their children into Grade 1. That is where the education management is needed. In order to get all schools into same level, first thing is to identify the schools which have lower performance in their academic. This requirement is satisfied in this study easily. So responsible people in education can introduce new methods to improve the academic performance in those schools.

As a result of free education, Sri Lanka obtained 90% literacy rate. In past decades, education system in Sri Lanka produced many scholars, scientists and administrators when comparing to recent decades. This is because, there will be some serious problems which may include in current education system. There is a decline in quality and standard of education at all levels. Current Education system is unable to provide physical infrastructure and resources for each and every school equally. Hence as a developing country reforming the education system is important and urgent priority. In order to reform the education system, first thing is to identify schools with same academic performance in primary education. So this study may helpful to identify the schools with same academic performance. By identifying schools which belong to same category, correction actions can be carried out easily. That means to share resources, knowledge and information with each other. This must be a great opportunity for all the children. Not only for the selected group of children. Schools which are situated in urban areas have more resources than schools in rural areas. Hence it is important to provide an equal opportunity to every child; otherwise it is against with principles of social justice and equity. In other hand it is an education management.

II. METHODOLOGY AND EXPERIMENTAL DESIGN

Cluster analysis can be viewed as dividing similar objects or data into categories or groups (clusters) that are meaningful, useful or both. When it comes to design meaningful clusters, natural structure of data are considered. If there is a large dataset, classification is a convenient method for summarizing those data in effective way. To generate test hypothesis depending on cluster structure is another reason for conducting cluster analysis.

There are various methods to group those objects into groups. These methods can be used to express similarity and dissimilarity of the objects. Measure of similarity depends on the nature of the variables or objectives of research. In clustering, it can be used for categorizing the objects.

Euclidean distance, Squared Euclidean distance, City-block (Manhattan) distance and Mahalanobis distance are some of similarity measures which are used in cluster analysis.

Hierarchical method is one way of conducting cluster analysis. Hierarchical clustering can be generalized as a series of successive merges or series of successive divisions. Agglomerative method and divisive method are two form of hierarchical clustering. In agglomerative method, initially there are many clusters because clustering starts with individual objects. That means initially objects are considered as clusters. Then the most similar objects are grouped into one cluster. Based on the similarity those groups are merged into one group. Single linkage, complete linkage, average linkage and wards method are some of the agglomerative methods that are used in this study. Divisive method is opposite method of agglomerative method. That means initially there is one group of objects. Then it divided into two sub groups based on dissimilarity. Objects in one sub group are far away from other sub group. In this method, those initial sub groups are divided further until one object remained.
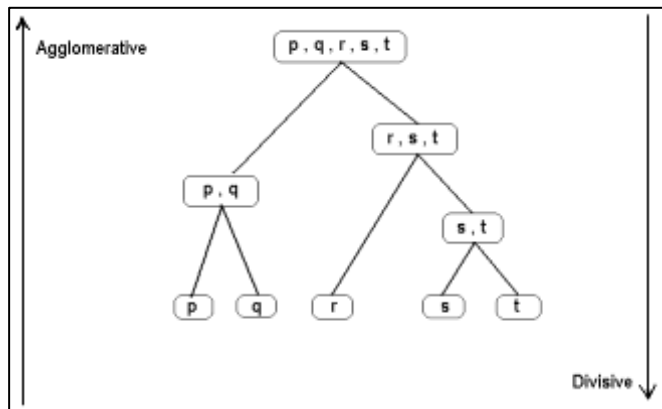

Figure1. Agglomerative and Divisive method

Data Manipulating
Dataset is obtained from the Department of Examination, Ministry of education. Data set contains mark of each student who faced to the scholarship examination in year 2013 & 2014. This raw mark of the Grade-v examination is used to cluster the schools. Schools name, type of school, medium of the school, zone name and the raw mark are variables in this data set. In this study, only Gampaha district schools (466) are considered in order to carry out the analysis. There are four zones included in Gampaha district. They are Minuwangoda (140), Gampaha (124), Kelaniya (102) and Negombo (100). Raw marks of each and every

student in those schools who sat Grade-v scholarship examination in year 2013 and year 2014 are used to carry out this analysis. In year 2014, 37245 students' records (raw mark) and 35227 students records (raw mark) in 2013 are used in this study. Furthermore this Grade-v scholarship examination results vary in between 0 and 200. These raw marks were used in this analysis after data manipulation

Data manipulation is a process of re-arranging data/ changing data in order to make it easier to read or more organized. When considering the above variables, only the raw mark can be used to carry on the clustering. Since data set contains only the raw mark, it is needed to create new summary record for each school in Gampaha district using these raw marks.

According to the rules and regulation in Sri Lanka, Grade-V scholarship marks are given out of 200. As per literature, there was concept called data pre-processing for clustering. Data pre-processing can be form of normalization/scaling, transformation/projection, cleaning, discretization and data reduction (Hichem, 2004). Discretization means divide the range of a continuous attribute into intervals. It can be known that in the best discretization, the interval cuts according to the context of the study and then it can make sense of the transformed attributes. This mark range (0-200) is divided into five ranges and created new variables by calculating the proportions. Though this method is subjective, dividing this range into five is done in different ways and chose most appropriate method. One method of dividing ranges as follows.

- Proportions of students who obtained marks between 0 and 70. (not qualified)
- Proportions of students who obtained marks between 70 and 100. (Average)
- Proportions of students who obtained marks between 100 and 160. (Pass)
- Proportions of students who obtained marks between 160 and 180.(cut off)
- Proportions of students who obtained marks between 180 and 200. (superior)

Reasons for selecting those margins to calculate above proportions:

- 70 or above mark; Department of Examination offer certificate for those students.
- 100 mark; half of the total mark of paper.

- 150-160 mark; in each year pass mark vary in between this range.
- 160-180 mark; To enter new school, mark should be in between this range.
- More than 180 mark; To obtain island rank, mark should be above 180.

III. RESULTS

Descriptive analysis is carried out for the data set in order to get an idea about the data set. Grade-V scholarship examination is held using two main languages, Sinhala and Tamil. According to the analysis 95.06% of schools in Gampaha district are Sinhala medium. That means students in those schools are sat to Grade-V scholarship examination using Sinhala medium. As mentioned before, each and every school in Sri Lanka is classified into four major categories namely type 1AB, type 1C, type 2 and type 3.
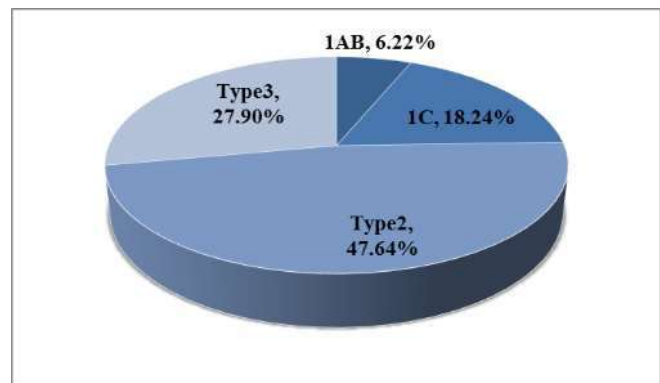

Figure2. General Classification of Schools

Most of schools are type 2 in Gampaha district indicating that schools having classes up to 11. As a percentage it is 47.64%. Fewer amounts of schools belong to type 1AB. As percentage it is 6.22%. And also 27.9% and 18.24% of schools belong to type 3 and type 1C respectively. In each district there are boys' schools, girls' schools and mixed schools. About 95% of mixed schools are situated in Gampaha district.

This analysis is carried out zone wise separately. According to the results obtained, most of schools are situated in the Minuwangoda zone. As a percentage, it is 30.04%. Nearly same numbers of schools are situated in Negombo and Kelaniya zone.
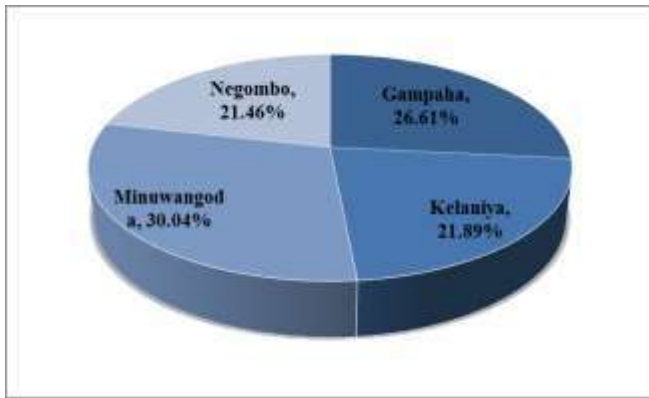
Figure3. Percentage of Schools in four zones

In each zone, Sinhala medium schools are higher than the Tamil medium schools. When considering the zone wise separately Minuwangoda zone has higher number of Sinhala medium school than other zones as well as it hasthe minimum number of Tamil schools among four zones. However each zone has Tamil medium schools. When comparing all types of schools, large number of type 2 schools (schools having classes with 1-11) are situated in each zone. Except Kelaniya zone, other zones have an order of having schools based on type. Gampaha, Minuwangoda and Negombo zones has higher number of type 2 schools and then type3, type 1C, type 1AB respectively. But in Kelaniya zone, type 1C schools are higher than the type3 schools.
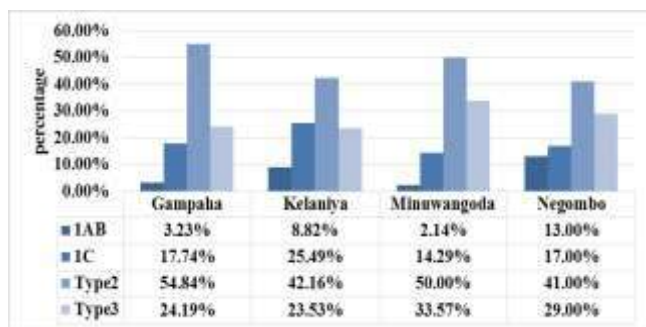

Figure4. Zones Wise Comparison Based on General School Classification

Furthermore it shows that that Gampaha zone has large number of type 2 schools while Minuwangoda zone has minimum number of type1AB schools. Large numbers of mixed schools and minimum number of boys schools are situated in each zone.

Advanced analysis is carried out using hierarchical clustering method. Hierarchical clustering method is used to figure out how schools are classified into different clusters in order to achieve the objectives. Wards method, single linkage, average linkage and complete linkage are four different techniques of agglomerative clustering method. Those four different techniques are carried out in each zone separately. PROC CLUSTER procedure in SAS statistical software is used to obtain following results. Hierarchical clustering method is implemented for each year separately. Cluster names are decided and identified by considering five variables in original data set and based on objects that are classified into groups. In this analysis dendrograms are used to identify the schools that belongs to each groups.
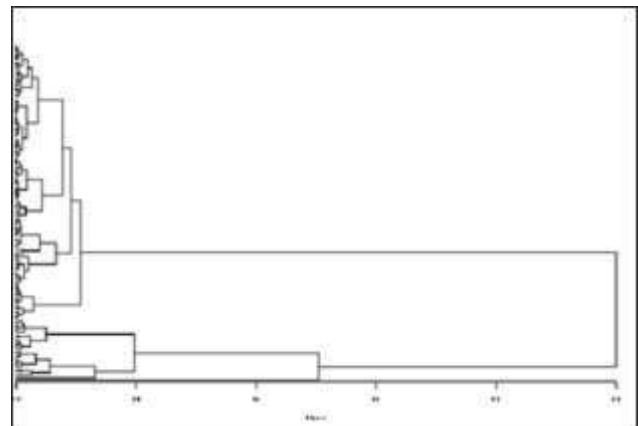

Figure5. Dendrogram for Ward's method under boys and mixed schools criteria (Year 2014)

Above dendrogram (Figure 5) is obtained using Ward's method. As previously mentioned in chapter 3, Ward's method belongs to one of the agglomerative hierarchical clustering procedure which minimizes total within cluster variation. By eye inspection, there would be 2 or 3 clusters. Anyhow it doesn't only depend on the dendrogram. Possible groupings depend on the objective of the researcher.
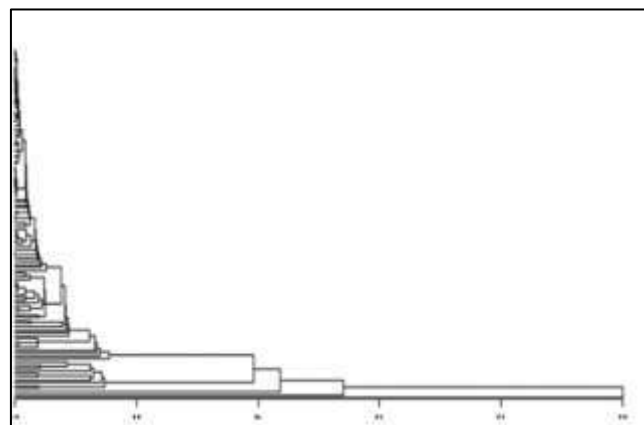
Figure 6. Dendrogram for single linkage method under boys and mixed schools criteria (Year 2014)

Above figure 6 is obtained from single linkage method. Here also all the schools under this criterion (boys and mixed schools) are categorized into clusters. But when comparing with Ward's method, way of clustering in here is not much clear than the Ward's method.
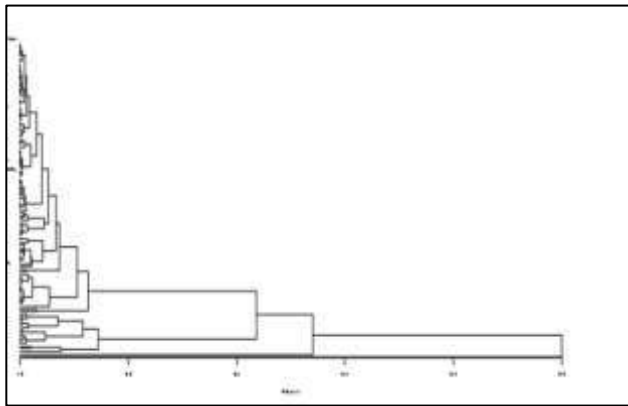

Figure 7. Dendrogram for average linkage method under boys and mixed schools criteria (Year 2014)

Average linkage method is also one of the agglomerative hierarchical methods that used here to obtain above figure 7. However this diagram illustrates some kind of clustering, but not like in Ward's method
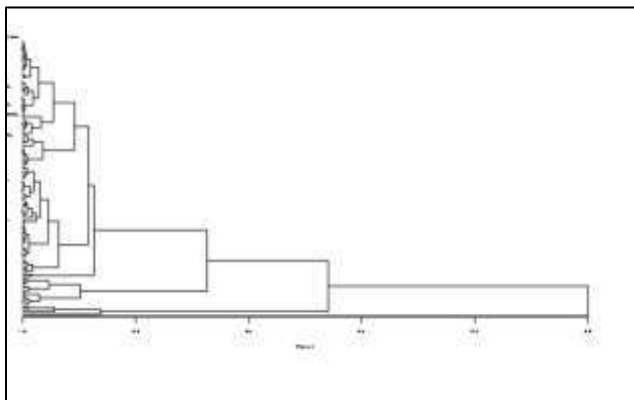

Figure 8. Dendrogram for complete linkage method under boys and mixed schools criteria (Year 2014)

Figure 8 is obtained from complete linkage method. However according to the dendrogram, it shows that schools are clustered somewhat better than single linkage and average linkage method.

The above dendrograms are used to figure out the cluster for each school in Gampaha zone by considering the all four hierarchical methods. Main difference between this classification method and the method currently existing in Sri Lankan education system is, this proposed method depends on the result of grade-v scholarship examination while currently existing method depends on A/L streams including in the particular school.

IV. DISCUSSION AND CONCLUSION

School clustering means grouping schools which have the same academic performance. One of the main approaches of identifying schools with similar performance is sharing facilities, resources and knowledge. Hence it can be helpful to promote collaboration, reflection, sharing and learning. In Sri Lankan education system, Schools are classified into four types depending on the classes within schools. That classification scheme is not based on the performance of schools. This study was carried out to identify performance of schools in Gampaha district schools based on Grade-v scholarship examination results and build a classification scheme.

Hierarchical method was used in this study to identify the performances of Gampaha district schools based on Grade-v scholarship examination result. Hierarchical methods such as wards method, single linkage, complete linkage and average linkage method were tried out in study. As per dendrograms it was found out that better separation of schools shows in wards method. Complete linkage was also better than single linkage and average linkage. However by eye inspection and depending on the objectives of research, it is suggested that creating three clusters namely Best, Moderate and lower.

All the four hierarchical clustering methods were used here to cluster the schools according to the criteria. According to the results obtained from the advanced analysis, there were schools which categorized into same group even if the hierarchical method changed. But some schools were classified into different groups. According to the result, same set of schools in Gampaha, Kelaniya, Negombo and Minuwangoda zone were identified as best schools and low performance schools in both years. However some of the schools belongs to the Moderate group was moved to the Best and Low performance groups in two years. By

13

considering all those results, conclusion concluded as follows.

- For each zones there are set of schools which have better performance of their academic in each method.
- Performance of most of schools can be identified using this classification scheme.

References

Doborawila S, Samita S, Thattil R*, et all*(2002). CLASSIFICATION OF STATE SCHOOLS IN SRI LANKA: A MULTIVARIATE APPROACH.*Ceylon Journal of Science: Physical Sciences, 9(I), 102-109*

Lock A,(2011) Clustering together to advance school improvement: working together together in peer support with an external colleague. *National College for Leadership of Schools and Children's Services.*

Bray M,*(1987)*Schools Clusters in Third World.*unesco-unicef co-Operative programme*

Giordano EA,2008)Schools cluster and teacher resource centers.*Paris: UNESCO:International Institute for Educational Planning*.

Pellini A, Bredenberg K, *et all*(2015). Basic Education Clusters in Cambodia. *Journal in Development in Practise*.

Nghatanga P,(2010). The role of cluster centre principals in the Ohangwena, Education region in Namibia, *Master of Education, Educational Leadership and Management*

Hichem F,(2004) Pre-Processing for Data Clustering.