

# A Cluster-Based Method for Evaluation of Truck's Weighing Control Stations

Abbas Mahmoudabadi<sup>1</sup> and Arezoo Abolghasem<sup>2</sup>

*<sup>1</sup>PhD Candidate, Technical and Engineering Faculty, Payam-e-Noor University & General Director of Traffic Safety Department, Road Maintenance and Transportation Organization, Tehran*

*<sup>2</sup>Transportation Engineer, Road Maintenance and Transportation Organization, Tehran Iran*

## 1. Introduction

Effective and safe road network is essential to economic growth and social welfare (Islamic Republic of Iran Majlis, 2005). More than 550 million tons of goods are transported by road network in Islamic Republic of Iran, which is a large amount of total cargo transportation's rate (RMTO, 2009). Overloading of vehicles increases safety risk and pavement damages (Planning and Management Organization, 1996), so to improve road safety and restrain pavement damages, legislation based on maximum allowable axle and total load is applied for heavy vehicles, which are passing on the roads (Mohammadi, 1999). Law enforcement control is done by weighing control equipments, which have been installed near the arterial roads, and to evaluate law enforcement, clustering method can be used as an effective technique.

Clustering method is used to classify data and decision making in transportation studies. Some researchers (Ahmed and Kanhere, 2007) used clustering method for public transport modeling in order to improve situation and possibility of changing decision in public transportation. They showed that using a step forward algorithm for clustering the nodes in finding routes problems while using clustering method, improves effectiveness of algorithm.

Clustering has been used to classify accident types (Depaire et al. 2008). They used a model based on clustering data, which divide accident data to 7 homogenous groups and finally analyzed injuries in each group. Comparing results based on divided groups and total data analysis showed that classifying data could be used for safety issues researchers with more accuracy.

(Fielding et. al 1985) used clustering method for analyzing transit systems equipments, considering vehicle's size, average speed, urban transportation situation, and bus specified routes, and have classified transit situation to 12 categories. Difference of characteristics for all categories was identified using variance analysis, detached group and tree decision-making method. Grouping data were made by seven main transit efficiency factors and finally some suggestion for each group was supposed to productivity evaluation.

In this paper after a brief review on methods of controlling cargo transportation and existing regulation in rural roads network, an evaluation method of trucks weighing control is presented using clustering analysis method. For evaluation and comparing results, available data on trucks weighing control, which have been collected by local authorities in all over

the country, have been analyzed based on clustering and finally weighing stations have been categorized in different groups.

## 2. Goods transportation controlling on road network

Overloading vehicles are important from two sides of decreasing transportation safety and fateful damages on roads surface and infrastructures. Overloaded vehicles in addition to damage road surface, because of Pre-designing of braking system particularly at slides and curves, endanger other vehicles. Increasing goods' transportation will lead to economical growth from one side of the view, and the other side will increase maintenance costs and safety risks. Studies show that the most important factor of roads destruction is heavy vehicle's axle load. Based on researches the damage on roads surface is related to the vehicle axle load by a non-linear acceleration rate mostly in a polynomial equation of forth degree. Because of limited budget in roads maintenance activities, it is essential to apply over loaded vehicles limitation rules effectively. In spite of heavy maintenance costs, do not repairing road surfaces leads to destroy vehicles and decreasing traffic speed, so pavement management is very important issue because of overloading vehicles passage .

At the present, the method of trucks weighing control in Iran is using weighing control equipments including static, dynamic, and portable scales at weighing stations (and Weigh in Motion System in 6 points). Maximum permitted axle load for different types of heavy vehicles, which has been identified by laws, is under the control of weighing stations.

## 3. Proposed method

According to importance of controlling overloaded vehicles, performance evaluation of weighing control is inevitable. Identifying effective parameters on weighing control, data standardizing and ranking weighing stations have done in this paper in the sequence process using a well-known method of statistical process of clustering.

### 3.1 Clustering method

Clustering method is a statistical method that divides observation data according to their similarity to homogenous and detached groups (Sharma, 1996). In this method, similarity criterion is firstly identified. The number of groups would be analyzed and recognized. Observation data in each group are similar to each other based on similarity criterion and each group is different from the other groups. It is possible to consider more than one variable simultaneously as similarity criterion in clustering method. However, there are some different methods of identifying similarity criterions, but the common method for defining similarity is known as equation number 1 while the similarity of  $i$ th and  $j$ th observed data of  $k$ th variable is defined by  $P_{ij}^2$ .

$$P_{ij}^2 = \sum_{k=1}^p (x_{ik} - x_{jk})^2 \quad (1)$$

Clustering is a hierarchically process means that the nearest observations are merged continuously to make larger groups. This process will be continued until number of groups is reached to the proposed number of desired clusters. Merging observations is done by different methods such as Centroid, Farthest- Neighbour, Nearest- Neighbour, Average

Linkage, and Ward method usually chosen by analysts based on acceptable frequency. Based on experience, experts' view and comparing results of clustering Ward method is used to merge data in this research.

### 3.2 Data normalization

Because of existing the effective parameters in decision-making techniques, if there are different dimensions, for variables normalization or standardizing methods are used to assimilate (Sharma, 1996). If data can be fitted by a normal distribution function, normalization method is used. In addition, there are different methods of output and input data normalization, which the most regular methods convert data to, numbers between 0 to 1 and -1 to 1. Equations (2) and (3) converts data to zero to 1 numbers and are used for data more than zero.

$$X_{new} = \frac{X_{old}}{X_{max}} \quad (2)$$

$$X_{new} = \frac{X_{old} - X_{min}}{X_{max} - X_{min}} \quad (3)$$

If it is necessary to convert data in the range of -1 to 1, equation (4) is used:

$$\left\{ \begin{array}{l} X_{new} = \frac{X_{old}}{X_{max}} \quad \forall X_{old} \geq 0 \\ X_{new} = \frac{X_{old}}{X_{min}} \quad \forall X_{old} \leq 0 \end{array} \right\} \quad (4)$$

In order to standardize data assuming distribution function is normal, equation (5) is used which  $X_{old}$  is preliminary amount of factor and  $X_{new}$  is standardized form:

$$X_{new} = \frac{X_{old} - \bar{X}}{\sqrt{\frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n-1}}} \quad (5)$$

### 3.3 Data gathering

To evaluate truck's weighing control performance, number of registered overloaded vehicles, amount of detected overload, the type of scale, average heavy vehicle's daily traffic, and number of overloaded vehicles with more than 5 tons overload are collected from all over the country in 6 month period to use defined below parameters:

- Number of registered overloaded vehicles of six months period
- Detected amount of overloading of six months period
- Proportion of registered overloaded vehicles of six months period to the same period in last year
- Proportion of detected amount of overloading of six months period to the same period in last year
- Average heavy vehicles daily traffic of six months period

- The fraction of number of overloaded vehicles with more than 5 ton overload to total registered overloaded vehicles of six months period

Tables 1 and 2 show maximum, minimum, average, and standard deviation of collected data before and after standardizing data respectively. As it is observed, to homogenize data, standardization is used because of having a normal distribution function.

Row	Parameter's Name	Experimental Data			
		Max	Min	Average	Standard Deviation
1	Number of registered overloaded vehicles	1754	2	193	255
2	Detected amount of overloading(ton)	4014	2	519	655
3	Proportion of detected overloaded vehicles of six months period to the same period in last year	615	-96	88	143
4	Amount of detected overloading proportion of six months period to the same period in last year	718	-94	89	162
5	Average daily heavy vehicles traffic	41059	103	3887	6493
6	Number of more than 5 ton overloaded vehicles to number of total overload proportion	0.35	0	0.07	0.09

Table 1. Parameters, experimental data

Row	Parameter's Name	Standardized Data			
		Max	Min	Average	Standard Deviation
1	Number of detected overloaded vehicles	6.12	-0.75	0	1
2	Detected amount of overloading(ton)	5.37	-0.79	0	1
3	Proportion of detected overloaded vehicles of six months period to the same period in last year	3.69	-1.29	0	1
4	Amount of detected overloading proportion of six months period to the same period in last year	3.89	-1.13	0	1
5	Average daily heavy vehicles traffic	5.72	-0.58	0	1
6	Number of more than 5 ton overloaded vehicles to number of total overload proportion	3.16	-0.87	0	1

Table 2. Parameters, standardized data

### 3.4 Setting number of clusters

The number of clusters is specified in two ways. In the first method, analysts specify number of clusters when the number of clusters needs to be applied to other decision-making process, for example using clustering method to fuzzy modeling. Second method of specifying number of clusters, is minimizing Within-Group Sum of Squares (Sharma, 1996) while the sum of total deviation is calculated and number of clusters is identified in such a way that inter clusters sum of squares, which is named missed sum of squares, would be minimized. Also proportional sum of deviation squares as Between-Group Sum of Squares is used as a criterion to specifying number of clusters.

Since the number of clusters in this research work is not too important, second method which is more scientific has been used and the number of clusters is specified in such a way that minimizes Within-Group, deviation. Table 3 shows total deviation, Within-Group and Between-Group deviation and proportion of Within-Group to total deviation which is decision making parameter.

According to table (3), if it is proposed that 85 percent of deviation is reserved, number of clusters would be 16, and if it is proposed to reserve 90 percent of deviation, number of clusters would be 22. As it is shown at table (3) rest of deviations after 22 would change slowly means increasing the number of clusters doesn't have significant rule in decision process.

Figure (1) shows changes in total deviation, Within-Group and Between-Group deviation based on the number of clusters. As it has been shown in this figure, difference between Total deviation and Between-Group deviation would be decreased after 16, since 16 covers 85 percent of deviations, so in this research 16 number of clusters has been considered to evaluate weighing stations performance in controlling goods transported by heavy vehicles.

Row	Number of Clusters	Deviation From Average Sum of Square Errors			
		Total	Within Groups	Between Groups	Proportion of Between Groups to Total
1	4	750	352	398	53
2	6	750	236	514	69
3	8	750	183	567	76
4	10	750	156	594	79
5	12	750	145	605	81
6	14	750	122	628	84
7	16	750	113	637	85
8	18	750	102	648	86
9	20	750	81	669	89
10	22	750	77	673	90
11	24	750	68	682	91
12	26	750	63	687	92
13	28	750	61	689	92
14	30	750	55	695	93

Table 3. Deviation from average sum of square errors

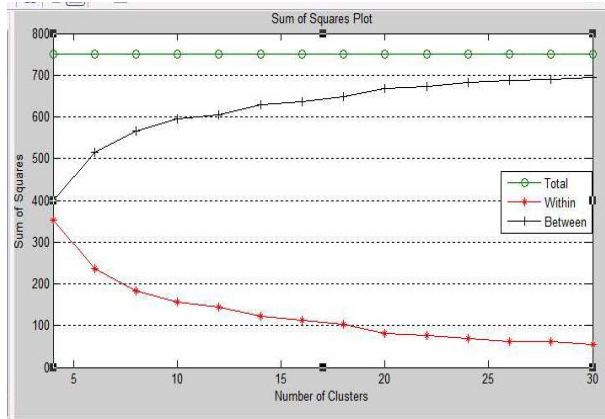


Fig. 1. Plot of square errors (Within, Between and Total)

#### 4. How to use MATLAB software

There is a command in MATLAB software, which has ability to cluster observations based on many parameters. As shown in tables 1 and 2, there are six parameters for clustering. Clusterdata is a well-known command in order to cluster data. There is a matrix 126×6 contains data used in this research work. If matrix named as X figure 2 shows the format of data.

Code	X(:,1)	X(:,2)	X(:,3)	X(:,4)	X(:,5)	X(:,6)
5707	0.011113	-0.0353	-0.64109	-0.81193	-0.46864	-0.86858
5704	0.819076	0.992241	-0.65484	-0.90436	-0.36345	-0.86858
...	...	...	...	...	...	...
5108	0.611202	1.519068	-0.08919	-0.82648	0.333153	0.958812
5106	-0.68703	-0.72033	0.783371	1.209399	0.448663	-0.86858

Fig. 2. Matrix contains data with six parameters

When command "clusterdata" is typed, results, are generated and are shown in a matrix contain two columns. The first column corresponds the row of observation and the second corresponds assigned cluster. For example if the first row of result matrix contains, 1 and 15, it means that station coded as 5707 in the first row of matrix X, belong to the 14th cluster.

The other command in MATLAB is kmeans, which is used to calculate the distance between observation and centre of the cluster. There is a output variable of SUMD in kmeans command to set distances and within group sum of square errors will be calculated by sum of the square distances of variable SUMD. Between sum of square errors will be calculated by total sum of square errors minus within sum of square errors while total sum of square errors is calculated as below (Sharma, 1996):

$$\text{Total sum of square errors} = (126-1) \times 6 = 750$$

When 126 is the number of observations and 6 is the number of standardized parameters.

## 5. Results

Considering the number of clusters by 16 and Ward Method to merging data, using clustering method of well-known software of MATLAB version R2008a, weighing station grades considering six identified variables are presented as table 4.

Class	Station's code	Number of stations
1	1126, 3603	2
2	3607, 3601	2
3	4507, 4102, 8708, 2109	4
4	1106	1
5	1125, 1104, 5408, 8701	4
6	5409, 8705, 8706, 3202, 2110, 9101	6
7	5108, 5107, 5109, 9703, 9702	5
8	1190, 1112, 3609, 3606, 3602, 7504	6
9	1108, 3608, 3604	3
10	9748, 4502, 8709, 3203, 3303, 7702, 2107	7
11	5102, 1122, 7503, 5401, 5406, 1403, 4110, 2104, 5703, 5704	11
12	5106, 1116, 4508, 4553, 4506, 4106, 41014, 4108, 4106, 4104, 6108, 6107, 2116, 2106, 9104, 9105	14
13	2105, 6401, 5402, 5403, 5404, 9701, 9506, 9502, 2111, 2105	10
14	5104, 1103, 1114, 7502, 4503, 4501, 4504, 1405, 4117, 6110, 6105, 8702, 7701, 2108, 2101, 5701, 5702, 5707	19
15	6402, 6404, 6403, 5407, 6106, 6102, 6101, 8704, 3201, 3301, 7706, 9505, 9501, 2103, 2102, 9103, 9102	17
16	3610, 7505, 6405, 4114, 4111, 4122, 4109, 4105, 4112, 4101, 6108, 6103, 7704, 9503, 5706	15

Table 4. Results (Clusters contain the number of stations)

## 6. Conclusion and future research

In this research, number of detected overloaded vehicles, amount of overload, scale's type, average heavy vehicle's daily traffic, and number of overloaded vehicles with more than 5 ton overload are collected from all over the country in a six month period and using standardization method are homogenized. By using minimum deviation method of Average within-Groups, number of clusters is calculated and considering Ward method for merging data, and using clustering method, the performance of weighing stations have been classified in 16 groups. Results show that this method is appropriate to use evaluation of weighing control performance in weighing stations.

Future studies, recommended using ranking method by similarity factors for traffic laws enforcement control, and any other important factors such as vehicles speed, overtaking, and left diversion control. Ranking enforcement stations and at a wider view ranking provinces could provide a comprehensive and precise image for managers about enforcement stations performance.

## 7. References

- Ahmed, S.; Kanhere, S.S, (Oct 2007), "Cluster-based Forwarding in Delay Tolerant Public Transport, Networks", 32nd IEEE Conference, Pages 625-634.
- Bolghari, M., (2010, 1389 local calender), "Driving laws and regulations", Second version, Department of transportation, Tehran municipality, 2010.
- Depaire, B., Wets, G., Vanhoof, K., (July 2008), "Traffic accident segmentation by means of latent class clustering", *Accident Analysis and Prevention*, Vol, 40, Issue 4, Pages 1257-1266.
- Fielding, G.J., Brenner, M.E., Faust, K., (May 1985), "Typology for Bus Transit", *Transportation Research, Part A*, Vol 40, Issue 4, Pages 1257-1266.
- Islamic Republic of Iran Majlis, (2005, 1384 local calender) "Forth five-year developing plan of Islamic republic of Iran", available on [www.majlis.ir](http://www.majlis.ir)
- Mohammadi Asa, A., (1999, 1378 local calender), "Laws and regulations in road transport", Road maintenance and transportation organization.
- Planning and management organization, (1996, 1375 local calender) "Road pavement and asphalt manual", Vol. 234.
- Road Maintenance and Transportation Organization, (2009, 1388 local calender) "Annual survey of road transport in Iran", available on [www.rmto.ir](http://www.rmto.ir)
- Subhash Sharma, (1996), "Applied Multivariable Techniques", University of South Carolina, USA.





## **MATLAB - A Ubiquitous Tool for the Practical Engineer**

Edited by Prof. Clara Ionescu

ISBN 978-953-307-907-3

Hard cover, 564 pages

**Publisher** InTech

**Published online** 13, October, 2011

**Published in print edition** October, 2011

A well-known statement says that the PID controller is the “bread and butter” of the control engineer. This is indeed true, from a scientific standpoint. However, nowadays, in the era of computer science, when the paper and pencil have been replaced by the keyboard and the display of computers, one may equally say that MATLAB is the “bread” in the above statement. MATLAB has become a de facto tool for the modern system engineer. This book is written for both engineering students, as well as for practicing engineers. The wide range of applications in which MATLAB is the working framework, shows that it is a powerful, comprehensive and easy-to-use environment for performing technical computations. The book includes various excellent applications in which MATLAB is employed: from pure algebraic computations to data acquisition in real-life experiments, from control strategies to image processing algorithms, from graphical user interface design for educational purposes to Simulink embedded systems.

### **How to reference**

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Abbas Mahmoudabadi and Arezoo Abolghasem (2011). A Cluster-Based Method for Evaluation of Truck's Weighing Control Stations, MATLAB - A Ubiquitous Tool for the Practical Engineer, Prof. Clara Ionescu (Ed.), ISBN: 978-953-307-907-3, InTech, Available from: <http://www.intechopen.com/books/matlab-a-ubiquitous-tool-for-the-practical-engineer/a-cluster-based-method-for-evaluation-of-truck-s-weighing-control-stations>

**INTECH**  
open science | open minds

### **InTech Europe**

University Campus STeP Ri  
Slavka Krautzeka 83/A  
51000 Rijeka, Croatia  
Phone: +385 (51) 770 447  
Fax: +385 (51) 686 166  
[www.intechopen.com](http://www.intechopen.com)

### **InTech China**

Unit 405, Office Block, Hotel Equatorial Shanghai  
No.65, Yan An Road (West), Shanghai, 200040, China  
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元  
Phone: +86-21-62489820  
Fax: +86-21-62489821

© 2011 The Author(s). Licensee IntechOpen. This is an open access article distributed under the terms of the [Creative Commons Attribution 3.0 License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.