

# A Choice of Representative Sections for the Estimation of the Transport Network Loading

Aleksey Vlasov<sup>1\*</sup>, and Ekatherina Gorshenina<sup>1</sup>

<sup>1</sup>Penza State University of Architecture and Construction, 28 Titova str., Penza, 440028, Russia

<sup>2</sup>Yuri Gagarin State Technical University of Saratov, 77 Politekhnicheskaya str., Saratov, 410054, Russia

**Abstract.** The problem of a choice of representative sections for transport network loading management is considered in the following article. To solve the problem, the use of Principal component analysis and the procedures of cluster analysis are offered. The example of the choice of representative sections at network traffic control is presented.

## 1 Introduction

Taking into account the quantitative growth of individual cars we can see that one of the characteristics of Russian cities are transport jams. They affect not only on availability of transport infrastructure, but also reduce efficiency of work of emergency services. In its turn it causes a potential threat of serious emergency situations and untimely help. Because of the bad organization of city transport system and appearance of jams a real economic damage exists. [7-10]

In the Russian Federation temporarily-dependent traffic light control with beforehand calculated signal schedules are traditionally used. [4-7,11]. In such situations when network loading comes nearer to the saturation state, any fluctuations in a traffic flow lead to spontaneous appearance of a transport jam. The key moment of solution of the given problem is the usage of modern methods of adaptive control of a city motion.

In most of known researches the attention is paid to computational methods of signal schedules, thus the attention is not practically paid to the supply with information of operation of control systems [9,12,13].

## 2 Problem statement

It is impossible to arrange with transport detectors all allowed directions of motion in the controlled area. Usually it is necessary to select representative subsets of points, information of which can give an objective estimate of a transport situation in all area of control [13-15].

There are, at least, three basic types of basic preconditions causing possibility of transition from the big number  $p$  of initial indicators of analyzed transport system condition to the essentially smallest number  $p'$  of the most informative sections. It is, first of all:

---

\* Corresponding author: [vlasov\\_a71@mail.ru](mailto:vlasov_a71@mail.ru)

- Duplication of information caused by strong correlation of intensity on adjacent sites;
- Small variability of received data for sections, intensity on which is poorly changed;
- Aggregation possibility (the simple or weighed summation) on some sections.

Technically, the task of passage with the least losses in self-descriptiveness to a new gang of sections of a network, in which measuring will be processed  $\tilde{z}^{(1)}, \tilde{z}^{(2)} \dots \tilde{z}^{(p)}$  can be presented as follows [2,3].

Let  $Z = Z(X)$  be some  $r$ -dimensional vector function of initial variables,  $x^{(1)}, x^{(2)} \dots x^{(p)}$  ( $p' < p$ ) and let  $I_{p'}(Z(X))$  be a definitely given measure of self-descriptiveness  $p'$ -dimensional system of indications  $Z(X) = (z^{(1)}(X), \dots, z^{(p)}(X))$ . Concrete select of a functional  $I_{p'}(Z(X))$  depends on specificity of the following actual task and is based on one of the possible criteria:

- The measure of auto self-descriptiveness aimed at maximum maintenance of information, containing in the initial array  $\{X_i\}_{i=1, \dots, n}$  concerning the initial indications;
- The measure of external self-descriptiveness aimed at maximum "squeezing" the information from  $\{X_i\}_{i=1, \dots, n}$ , containing in this array concerning some external indexes.

The task is in definition of such feature set  $\tilde{Z}$  discovered in class  $F$  of admissible conversions of initial indexes  $x^{(2)} \dots, x^{(p)}$ , that

$$I_{p'}(\tilde{Z}(X)) = \max_{Z \in F} \left\{ I_{p'}(Z(X)) \right\}. \quad (1)$$

Both of variants of a concrete definition of this setting (defining the concrete select of standard of self-descriptiveness  $I_{p'}(Z(X))$  and a class of admissible conversions) leads to the concrete method of lowering of dimensionality.

### 3 Task solution

The method of principal component analysis (PCA) belongs to the methods, which allow to reduce dimensionality. The method provides usage of various linear orthogonal normalized combinations of initial indexes as a class of admissible conversions  $F$ :

$$z^{(j)}(X) = c_{j1} \left( x^{(1)} - \mu^{(1)} \right) + \dots + c_{jp} \left( x^{(p)} - \mu^{(p)} \right); \quad (2)$$

$$\sum_{v=1}^p c_{jv}^2 = 1, j = 1, 2, \dots, p; \quad (3)$$

$$\sum_{v=1}^p c_{jv} c_{kv} = 0, j, k = 1, 2, \dots, p; j \neq k; \quad (4)$$

where  $\mu^{(v)} = Ex^{(v)}$  - a mathematical expectation  $x^{(v)}$ .

As a self-descriptiveness standard  $p'$ -dimensional system of indexes  $(z^{(1)}(X), \dots, z^{(p)}(X))$  the expression is used:

$$I_{p'}(Z(X)) = \frac{Dz^{(1)} + \dots + Dz^{(p)}}{Dx^{(1)} + \dots + Dx^{(p)}}, \quad (5)$$

where  $D$  is a character of dispersible evaluation.

Following the general optimization setting of the task of size lowering and supposing a parsed as  $X$   $r$ -dimensional aleatory variable measure with a vector of average values

$\mu = (\mu^{(1)}, \dots, \mu^{(p)})$  and covariance matrix  $\Sigma = (\sigma_{i,j})$  we can define a measure of self-descriptiveness  $I_{p'}(Z)$  by auxiliary  $p'$ -dimensional one of the system of indexes  $Z = (z^{(1)}, \dots, z^{(p')})$  by means of (2). Then at any fixed  $p' = 1, 2, \dots, p$  the vector of required auxiliaries  $\tilde{Z}(X) = (\tilde{z}^{(1)}(X), \dots, \tilde{z}^{(p')}(X))'$  is defined as a linear combination:

$$\tilde{Z} = LX, I_{p'}(\tilde{Z}(X)) = \max_{Z \in F} \{I_{p'}(Z(X))\}; \tag{6}$$

where  $L$  is a matrix and its lines satisfy to the orthogonality requirement:

$$L = \begin{pmatrix} l_{11} & \dots & l_{1p'} \\ \vdots & \ddots & \vdots \\ l_{p'1} & \dots & l_{p'p'} \end{pmatrix}. \tag{7}$$

Gained in this way variables  $\tilde{z}^{(1)}(X), \dots, \tilde{z}^{(p')}(X)$  are called as principal components of the vector  $X$ . The first principal component  $\tilde{z}^{(1)}(X)$  of a researched system of indexes  $X = (x^{(1)}, \dots, x^{(p)})$ , is a such normalized centered linear combination of these indexes, which among all other possesses the greatest variance, and the  $k$  principal component is a such normalized centered linear combination of indexes, which is not correlated with  $k-1$  previous principal components and, among all other, possesses the greatest variance.

### 4 Method application

Now we consider finding of an amount and a dislocation of controlled areas which can allow us to have the objective information on network loading. Thus, intensity of motion for cuts, where transport detectors were not set, can be defined as a function of received measuring.

In the table 1 field examinations outcomes of the traffic flow on Kirov street's intersections are presented (city of Penza).

**Table 1.** Traffic flow, vehicle./hour.

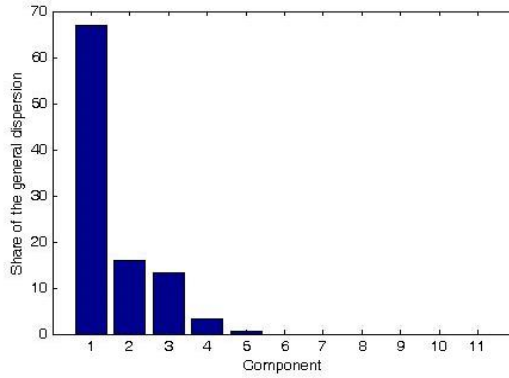
Direction /Time	1	2	3	4	5	6	7	8	9	10	11
7:00	144	180	300	96	192	168	276	144	360	180	300
9:00	420	324	468	336	252	528	456	420	720	348	372
12:00	468	600	432	336	216	468	564	540	696	312	600
15:00	360	516	228	240	156	432	384	264	624	252	612
18:00	420	492	564	276	132	252	528	504	516	216	516
21:00	276	312	468	144	180	372	324	360	432	120	396

For counting of areas of control in a transport network we take advantage of PCA-analysis function (on the given matrix of covariances) application package Matlab. [1-3].

Outcomes of evaluations are the percentage shares of the general variance, explained by principal components for the given observations of a traffic flow, represented in the figure 1. As we can see, for an explanation of 99.4196 % of the general variance in the considered network fragment there is enough usage of 4 representative cuts.

The concrete definition of representative cuts for installing transport detectors is fulfilled by clusterization of given measures and pointing out 4 clusters based on the amount of principal components (figure 2). As the near metric of objects the Euclidean distance was

used. The building-ups of the hierarchical tree of clusters were made on algorithm of a centroid, which uses the "barycentres" distance of groups. As we see, clusters №1 and 3 contain one direction of motion, clusters № 2 and 4 contain 6 and 3 units accordingly.



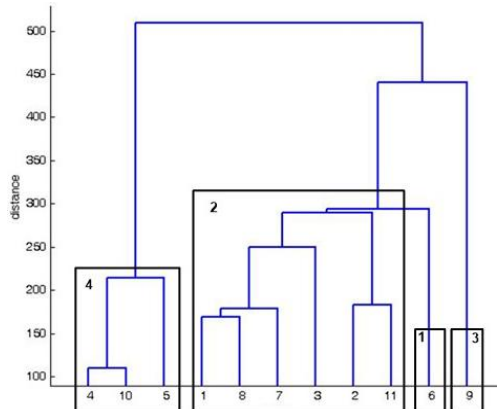
**Fig. 1.** The percentage shares of the general variance explained by principal components.

Let's consider procedure of select of representative cut for the cluster №2 in detail. Intensity of motion for cluster directions has similar dynamics (figure 3). As the cluster centre we select the direction having the minimum aggregate distance to other units of the cluster (table 2).

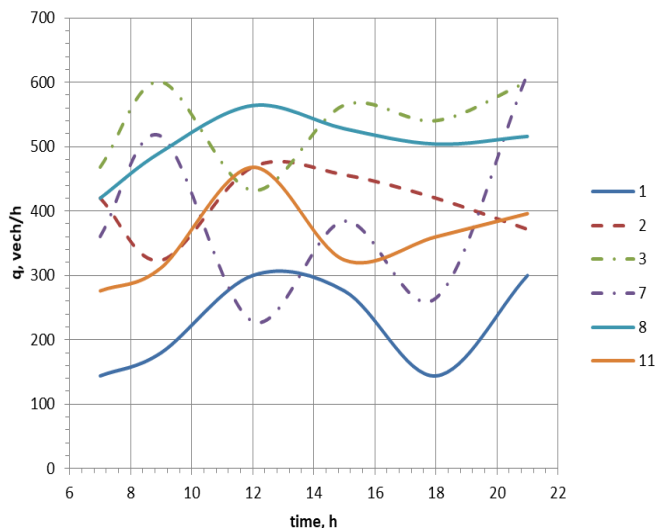
**Table 2.** Distance between units of a cluster №2.

Direction	1	2	3	7	8	11
1	0	242,39	320,87	206,11	168,85	362,19
2	242,39	0	419,31	216,33	282,96	183,17
3	320,87	419,31	0	253,99	234,23	438,61
7	206,11	216,33	253,99	0	188,59	257,37
8	168,85	282,96	234,23	188,59	0	390,87
11	362,19	183,17	438,61	257,37	390,87	0
The total	1300,41	1344,16	1667,01	1122,39	1265,5	1632,21

As we see from results of calculations for the cluster №2, it is expedient to select a direction №7 as the representative cut. The correlation analysis has shown high values of correlation coefficient of the traffic flow dynamics in a cluster with representative cut (table 3) that testifies the effectiveness of the offered method.



**Fig. 2.** Representation of a traffic flow in the form of a tree of clusters.



**Fig. 3.** Dynamics of traffic flow for a cluster №2.

**Table 3.** Correlation of a traffic flow directions with the representative cut.

Direction	1	2	3	7	8	11
7	144	180	300	276	144	300
9	420	324	468	456	420	372
12	468	600	432	564	540	600
15	360	516	228	384	264	612
18	420	492	564	528	504	516
21	276	312	468	324	360	396
Coefficient of correlation	0,93	0,80	0,87	1,00	0,92	0,89

## 5 Conclusion

The usage of combination PCA and cluster analysis procedures while allocating of transport detectors allows to reduce their amount considerably in cases of minimal informational losses about loading of a transport web. In this example, to provide a system for managing information about the network load, it is enough to install transport detectors in 4 of 11 directions, while the loss of information about the dispersion of traffic flows will not exceed 1%.

For development and perfection of the offered method additional examinations of cluster analysis procedure are required. The algorithm, used in examination, gives clusters, which are unequal in an amount of directions.

## References

1. V.P. Djakonov, *MATLAB 6.5 SPI/7 + Simulink 5/6 in the mathematician and modelling*. A series «Library of the professional». - Moscow: SOLON-PRESS (2005)
2. S.A. Ajvazjan, V.M. Buhstaber, I.S. Enjukov, L.D. Meshalkin *The applied statistics: Classifications and dimensionality lowering*. Moscow: the Finance and statistics (1989)
3. G.H. Dunteman, *Principal components analysis*. SAGE Publications, (1989)

4. E.B. Hilazhev, V.S. Sokolovsky, M. Gurulev, J.I. Zajdenberg, *Systems and tools of automated management by traffic in the cities* Moscow: Transport (1984)
5. N.O. Brajlovskij, B.I. Granovskij, *Traffic control of vehicles*. Moscow: Transport; (1975)
6. A. Vlasov, *Features of Calculation of Traffic Light Control Modes in the Conditions of Intensive Road Traffic*. 12th International Conference "Organization and Traffic Safety Management in large cities", SPbOTSIC-2016, 28-30 September 2016, St. Petersburg, Russia, Transportation Research Procedia vol. **20** (2017) p. 676 – 682 (2016)
7. A.A. Vlasov, *Urban adaptive traffic control*. Penza: Publishing House of the Penza State University of Architecture and Construction; (2012)
8. M.P. Pechersky, V.G. Khorovich, *Automated traffic control systems*, Moscow: Transport; (1979)
9. S. Lin, Y. Xi, *An efficient model for urban traffic network control*. Proceedings of the 17th World Congress. The International Federation of Automatic Control, pp. 14066-14071, Seoul, Korea; (2008)
10. Lagerev, R. Yu., Mikhaylov, A. Yu., Lagereva, S.V. *Methods of prevention of transport network*. Bulletin of the Scientific Center of Children's Safety №5. pp. 82-88 (2010)
11. V.T. Kapitanov, Ye.B. Khilazhev, *Management of urban traffic flows*. Moscow: Transport; (1985)
12. P.B. Hunt, D.I. Robertson, R.D. Bretherton, R.I. Winton *SCOOT - A Traffic Responsive Method of Coordinating Signals*, TRRL Report 1014. Transport and Road Research Laboratory, Crowthorne, Berkshire, UK; (1981)
13. C. Bielefeldt, F. Busch, *MOTION - a new on-line traffic signal network control system*, In proceedings 7th International Conference "Road Traffic Monitoring and Control", London, pp. 55-59; (1994)
14. J. Xuexiang, Yi. Zhang, Y. Danya *Simultaneously Prediction of Network Traffic Flow Based on PCA-SVR*. Advances in Neural Networks – ISNN 2007 4th International Symposium on Neural Networks, ISNN 2007, Nanjing, China, June 3-7, 2007, Proceedings, Part II pp 1022-1031 (2007)
15. B.L. Smith, B.M. Williams, R.K. Oswald, *Comparison of Parametric and nonparametric models for traffic flow forecasting*. Transportation Research C, **10**(4), 303-321. (2002)