The 9th International Conference on
Electrical and Control Technologies
May 8–9, 2014, Kaunas, Lithuania

ANALYSING VOTING BEHAVIOR OF THE LITHUANIAN PARLIAMENT USING CLUSTER ANALYSIS AND MULTIDIMENSIONAL SCALING: TECHNICAL ASPECTS

Vytautas MICKEVIČIUS*, Tomas KRILAVIČIUS*,**, Vaidas MORKEVIČIUS***
*Vytautas Magnus University, Lithuania; **Baltic Institute of Advanced Technology, Lithuania;
***Kaunas University of Technology, Lithuania

Abstract: Rational models of electoral behavior emphasize the need of sufficient information for voters to make their decisions. Monitoring the behavior of a single politician is not easy to implement, not to mention of the whole parliament, since for the latter one must apply statistical methods designed for the analysis of large amounts of information. In this paper we propose methods and techniques for the analysis of voting behavior of the Lithuanian Parliament (Seimas) that allow for clearer identification and recognition of voting patterns of the Seimas. Votes of the last sessions of the 2008-2012 term of the Seimas (pre-election period) are analyzed employing cluster analysis. Also, multidimensional scaling is used to visualize the generated results. Results obtained using different vote coding methods and clustering techniques are compared in the paper, too.

Keywords: clustering, multidimensional scaling, political science

1. Introduction

One of the main features that characterize democratic countries are free elections [1]. In Lithuania, there are four periodic elections taking place: President, Parliament (the Seimas), municipalities and European Parliament. The former two – President and Seimas – are considered the most important. The actions of both the President and the Seimas are being quite closely monitored during their terms. This is quite natural -- the aforementioned governmental bodies are formed by majority voting of Lithuanian citizens themselves, therefore, politicians are expected to serve the interests of the electorate. One of the methods to evaluate this is monitoring and analyzing behavior of the elected politicians.

It is possible to monitor the behavior of the President by directly analyzing his decisions and opinion on certain issues. However, the analysis of the parliamentary behavior is a much more challenging task since one has to monitor many people simultaneously (there are 141 members in the Seimas). Thus, it is crucial to apply appropriate research methods and techniques for the analysis of parliamentary behavior patterns. In this paper we direct our attention to the analysis of voting of members of the Seimas (MPs) with the aim to identify the homogeneity inside the factions, as well as their positions in regard to each other.

An assumption is made that a solid and disciplined faction is more influential than a faction of the same size but less disciplined. Accordingly, outcomes of voting are more strongly influenced by more disciplined factions. Factions of the Seimas are formed on the basis of the elected parties. They usually make coalitions, when different factions agree to work jointly. Factions forming the governing majority are usually called position, and factions opposing government's programme are called opposition. On most issues opinions of position and opposition are very different and this is well reflected in the voting behavior patterns of MPs [3], [6].

However, sometimes opinions are voiced that some MPs are voting differently from their factions while seeking personal benefits or acting on the behalf of certain interest groups. It is doubtful if this assumption could be true, yet testing it is a quite important and interesting task from the political science perspective.

While seeking to introduce proper methods and techniques for voting analysis of LR Seimas, the main tasks being tackled in this paper include:

1. Examining MPs loyalty towards their factions, as well as position or opposition;
2. Examining congruence between real composition of the factions and statistically defined groups based on the voting results;
3. Identifying MPs routinely voting differently from factions they represent, and similarly to MPs from other factions.

The analysis of parliamentary voting is a rather well researched subject [2], [3], [4]. Lithuania is no exception, i.e. work of the Seimas has been analyzed using various methods from both political science [5] as well as statistical perspectives. Political scientists and statisticians working together also produced interesting

results [6]. Importantly, quite many different methods has already been applied, such as homogeneity analysis [7], social networks analysis [8], multidimensional scaling [9].

Cluster analysis was also been applied for the analysis of the Lithuanian parliament voting [9]. The main challenges discerned include: selecting a proper coding for voting data, selecting the best dissimilarity measures, and selecting the best methods and techniques for analysis and visualization of the results.

2. Data

2.1. Seimas in the pre-electoral period

Roll-call voting (further on *voting* or *roll-call*) data used in the research was taken from the database of the project *atviras-seimas.info* [10]. Time period analyzed was the end of the 2008-2012 year term of LR Seimas (pre-election period for 2012-2016 year term). This period lasted from 2012-03-10 to 2012-11-16 and included 8th ordinary session and both 9th ordinary and 9th extraordinary sessions of LR Seimas. There existed 8 factions during the analyzed time period (Table 1).

Table 1. Factions in Seimas during pre-electoral period

Abbrev.	Full name of the faction	Political position
DPF	Labour party faction	Opposition
FTT	Faction „Order and justice“	Opposition
KPF	Christian party faction	Opposition
LCSF	Liberal and center union faction	Position
LSDPF	Lithuanian social democratic party faction	Opposition
LSF	Liberal movement faction	Position
TS-LKDF	Homeland union Lithuanian christian democrats faction	Position
Other	MPs who represented several factions or Mixed group during the period analyzed	Other

There were analyzed 1489 roll-call votes in total. Two MPs had to be removed from the original dataset as there were lack of information about the majority of their votings. Therefore votings of the 139 remaining MPs were analyzed.

2.2. Coding of votes

Every roll-call in Seimas can have 6 different outcomes:

1. **Aye**: voted for the bill or proposal;
2. **No**: voted against the bill or proposal;
3. **Abstain**: abstained during the voting;
4. **No vote**: registered for the voting but did not vote;
5. **No participation**: did not participate in the plenary sitting when the voting was taking place;
6. **No info**: there is no information about the outcome of the voting.

In order to analyze votings using statistical methods we must encode their outcomes with numeric values. There are quite a few different ways to do that, but in this paper we examine 2 particular codings at length (Table 2).

Table 2. Coding of votes

	Standard	Alternative
Aye	2	1
No participate	1	0
No vote	0	–1
Abstain	–1	–1
No	–2	–1
No info	–10	0

Standard coding uses different numeric value for each different outcome of the voting, hence in this case there is a difference how exactly an MP opposed the bill or proposal (abstained, did not vote or voted against). The goal of standard coding is to compare a position of MP with his factions position in the most accurate way (i.e. if the majority of the faction abstains, but one of its members votes against, his position is a little bit different than it "should be"). Numeric values used in the standard coding were introduced in the [6] paper. "No info" value was intentionally selected unlike the others in order for outliers (MPs with too much information of their votings missing) to be clearly seen among the other MPs.

The goal of the alternative coding is to present a simplified view of general tendencies. Keeping in mind that in most votings principle *not aye = no* is correct (the bill or proposal is passed according to the number of *ayes*), voting outcomes *no*, *abstain* and *no vote* are considered equal, meaning that MP opposes the bill or proposal. Absence of MP in the plenary sitting may have many reasons, therefore *no participation* is considered a neutral outcome, as well as *no info*.

3. Tools and methods

The research (both clustering and multidimensional scaling) was performed using R [11]. It is a free software programming language and software environment for statistical computing and graphics. The R language is widely used among statisticians and data miners for developing statistical software and data analysis [12]. R enables its users to build and develop their own additions to R, also known as packages [13].

Clustering, or cluster analysis is a distribution of objects into several groups according to their similarity [14]. In turn, cluster is a group of similar objects. The goal of clustering is to distribute the objects into clusters in such a way that the differences between objects inside the clusters would be as little as possible, while keeping the difference between clusters as big as possible.

In order to estimate the similarity of the objects a variety of **dissimilarity measures** are used. Most common dissimilarity measures are metric. A common name for metric dissimilarity measures is a distance [14]. Most popular distances are Minkowski distances.

$$\left(\sum_{i=1}^{m}|x_i - y_i|^l\right)^{1/l}, \quad l > 0. \tag{1}$$

Separate cases of Minkowski distance: Manhattan with l = 1); Euclidean with l = 2 and Chebyshev with l infinity

3.1. K-means method

One of the most popular clustering methods is *k*-means [15]. It is best used for large datasets. The following algorithm is being used while applying *k*-means:

1. Objects are distributed into *k* initial clusters;
2. Distance between the object and the center of every cluster is calculated for every object;
3. Objects are assigned to the nearest cluster;
4. Centers of every cluster are recalculated;
5. Steps 2-4 are repeated until there are no more redistribution of the objects.

An inconvenience may occur while applying *k*-means method – it is necessary to define the number of clusters beforehand. There is a wide variety of techniques for that [15], but in this paper we stick to 8 clusters because that as well is the number of factions.

3.2. Clustering quality evaluation

In this paper we apply 2 internal and 2 external clustering quality measures (also called validation indexes).

Dunn index. This criterion is internal, which means it defines the accuracy if the clustering process itself, regardless of the classes assigned to the objects a priori [17], see eq. (2).

$$DUNN = \min_{1\le i\le c}\left\{\min\left\{\frac{d(c_i,c_j)}{\max_{1\le i\le c}(d(X_k))}\right\}\right\}; \quad (2)$$

here $d(c_i,c_j)$ is a distance between clusters X_i and X_j, $d(X_k)$ defines distances between objects inside cluster X_k, and c is the number of clusters.

To put it simply, *Dunn* index is a ratio of the smallest distance between clusters, and largest distance between two objects in the same cluster. The aforementioned distance in this case is the shortest way between two points – an Euclidean distance.

Davies-Bouldin (DB) index. This is also an internal quality evaluation criterion [17], see eq. (3).

$$DB = \frac{1}{c}\sum_{i=1}^{c}\max_{i\ne j}\left\{\frac{d(X_i)+d(X_j)}{d(c_i,c_j)}\right\}; \quad (3)$$

here c is the number of clusters, $d(X_i)$ and $d(X_j)$ are distances between objects and centers of the clusters X_i and X_j respectively, and $d(c_i,c_j)$ are the distances between the aforementioned clusters.

Davies-Bouldin index is a similar measure to *Dunn* index eq. (5), however the values used in the latter are calculated generally for all clusters, while in *DB* index the values are calculated for every cluster separately and the quantity is then divided by the number of clusters (an average is produced).

Rand index. This criterion defines a similarity between statistically structured clusters and initial classes (factions in this case) [16]. It is an external quality evaluation criterion (Eq. 4), here *TP*, *TN*, *FP* and *FN* are defined in Table 3.

$$RI = \frac{TP+TN}{TP+TN+FP+FN}. \quad (4)$$

Table 3. Definitions of *TP*, *TN*, *FP* and *FN*

	Is in the class	Is not in the class
Assigned to cluster	*TP* – true positive	*FP* – false positive
Not assigned to cluster	*FN* – false negative	*TN* – true negative

Rand index is a ratio of correct decisions to all decisions made by clustering algorithm.

Purity index. This is also an external quality evaluation criteria. *Purity* is a simple and popular validation measure [16]. We get its value using formula:

$$PURITY(Q,Z) = \frac{1}{n}\sum_{i}\max_{j}(q_i\cap z_j); \quad (5)$$

here n is the number of objects, $Q=(q_1,q_2,\dots,q_i)$ is a set of clusters (i – number of clusters), $Z=(z_1,z_2,\dots,z_j)$ is a set of initial classes (j – number of initial classes).

Basically , the value of *Purity* index is produced by totaling the amount of objects of the most popular class in each cluster and dividing the count by the number of objects.

When evaluating quality of the clustering it is desirable to get a value of *Davies-Bouldin* index as little as possible, while the higher value 3 other criteria produces the better the clustering is.

3.3. Visualization of multidimensional data

Multidimensional data can be visualized by applying direct or projection methods. Direct methods has the advantage of maintaining 100% of explained variance, or in other words, no initial information is lost during the process of visualization. However, processing many variables by using direct methods often becomes a complicated task due to readability of the results – it is very difficult or nearly impossible to understand and interpret them [18]. The goal of projection methods is to reduce the amount of variables (dimensions) to 2 or 3 in order to clearly and understandably present them in a graphic form. Naturally, the consequence of such actions is a loss of a certain percentage of initial variance explained.

It is very important to choose proper amount of dimensions when using projection methods. For example, 4-dimensional plot would contain more initial information than a 3-dimensional one, yet it is the latter that would be much easier to understand and interpret. General rule of thumb states that we cannot lose a relatively large part of variance explained just for the sake of simplicity, and vice versa – we should prefer simplicity over a relatively small part of variance explained.

3.4. Multidimensional scaling (MDS)

Suppose I is a dataset (a set of objects). Let $\delta_{i,j}$ be the distance between objects i and j. All these distances together produce a square distance matrix:

$$\Delta = \begin{pmatrix} \delta_{1,1} & \delta_{1,2} & \dots & \delta_{1,I} \\ \delta_{2,1} & \delta_{2,2} & \dots & \delta_{2,I} \\ \vdots & \vdots & \ddots & \vdots \\ \delta_{I,1} & \delta_{I,2} & \dots & \delta_{I,I} \end{pmatrix}. \quad (6)$$

Having Δ matrix, the goal of multidimensional scaling [20] is to find I vectors $x_1, \dots, x_I \in \mathbb{R}^N$ such as the following condition would be met: $\|x_i - x_j\| \approx \delta_{i,j}$, for all $i, j \in I$. Here $\|\cdot\|$ is a vector norm. The most commonly used norm is an Euclidean norm (it is used in this paper as well).

Using the eigenvector of distance matrix Δ we can find the percentage of variance explained [19]:

$$VE = \frac{\sum_{i=1}^{m} \lambda_i}{\sum_{j=1}^{n} \lambda_j}; \quad (7)$$

here m is a chosen number of dimensions, n is a full number of dimensions, λ_i and λ_i are i-th and j-th eigenvalues respectively. To put it simply, $\sum_{i=1}^{m} \lambda_i$ is a sum of m first eigenvalues, and $\sum_{j=1}^{n} \lambda_j$ is a sum of all eigenvalues.

4. Experimental evaluation

4.1. Data preparation

The original data was stored in the CSV type files.

The initial data matrixes had voting results stored in their rows and the IDs of MPs in their columns. Hence these matrixes had to be transposed, turning MPs into objects (rows) and votings into variables (columns).

Two data files were prepared in total – for two different codings of votes (standard and alternative).

4.2. Clustering of the MPs

The number of clusters was chosen according to the number of initial classes (factions in this case) – 8. It is an optimal choice for comparing the results with initial data.

3 different distances were applied for each of the 2 different codings of votes, thus generating 6 sets of clustering results in total. Best versions of clustering were selected according to their quality evaluation (Table 4).

Table 4. Clustering quality evaluation

Coding	Distance	Rand	Purity	Dunn	DB
Standard	Manhattan	0,845	0,683	0,543	3,047
	Euclidean	0,823	0,676	0,469	2,956
	Chebyshev	0,823	0,676	0,469	2,956
Alternative	Manhattan	0,845	0,662	0,627	3,364
	Euclidean	0,788	0,583	0,614	3,27
	Chebyshev	0,788	0,583	0,614	3,27

Identical clustering results were produced when using Euclidean and Chebyshev distances (for both standard and alternative coding), therefore the values of their validation indexes are identical as well.

Considering the external quality evaluation criteria (Table 4, *Rand*, *Purity*) and assuming that clusters should resemble factions (initial classes) as accurately as possible, it can be proposed that the standard coding and Manhattan distance are a little bit better than the others.

To estimate quality of clustering internal validation indexes should be taken into account (Table 4, *Dunn*, *Davies-Bouldin*). In this case the results of alternative coding are visibly better than those of the standard one. However, it is quite difficult to define the best distance – *Dunn* index marginally favors Manhattan distance, and *DB* index suggests Euclidean or Chebyshev distance by the same narrow margin. Manhattan distance is examined further by choice.

Tables 5 and 6 displays the distribution of objects (MPs) into clusters and initial classes (factions). Numbers in bold indicate the cluster being assigned to the faction in the same row.

Table 5. Manhattan distance, standard coding

			Clusters							
			1	2	3	4	5	6	7	8
Factions	Opposition	DPF	–	–	6	–	2	1	1	–
		FTT	–	–	1	–	**11**	1	1	**3**
		KPF	–	–	–	1	–	6	–	–
		LSDPF	–	–	**15**	1	–	–	**6**	–
	Position	LSF	9	–	1	1	–	–	–	–
		LCSF	2	1	–	**8**	–	–	–	–
		TS-LKDF	**16**	**26**	–	1	–	–	–	–
		Other	–	1	3	2	–	**10**	2	–

Table 6. Manhattan distance, alternative coding

			Clusters							
			1	2	3	4	5	6	7	8
Factions	Opposition	DPF	–	–	6	–	2	1	1	–
		FTT	1	2	–	–	**9**	–	2	**3**
		KPF	–	1	–	–	–	–	6	–
		LSDPF	–	**15**	–	1	–	**5**	1	–
	Position	LSF	–	1	8	1	–	–	1	–
		LCSF	1	–	1	**9**	–	–	–	–
		TS-LKDF	**30**	–	**12**	1	–	–	–	–
		Other	2	1	–	2	1	2	**9**	1

42 out of 43 members of TS-LKDF (which is the core faction of position) were assigned to 2 clusters (26 in one, 16 in another). In the same way, 21 out of 22 members of LSDPF (which is the main faction in opposition) were assigned to 2 other clusters (15 and 6 respectively). These results point out that although the locations of these factions are rather clear, polarization do exist in the two largest factions of Seimas. Otherwise the vast majority of faction members would be assigned to one cluster.

In accordance with *Purity* index, we can state that the factions of position are more solid than those of the opposition (Table 7). It is determined by the lack of strong and clear location of the factions in opposition. Meanwhile, the location of LSF and LCSF appears to be very solid – that partially compensates the polarization in TS-LKDF.

Table 7. Manhattan distance

	No. of MPs	Purity criterion	
		Standard	Alternative
Position	65	0,769	0,785
Opposition	56	0,625	0,571
Other	18	0,556	0,5

4.3. Visualization and analysis of clustering results

Before we generate the results in a graphic form we must choose the number of dimensions that will contain the aforementioned results. The percentages of variance explained according to number of dimensions are displayed in Table 8.

Table 8. Variance explained

Coding	No. of dimensions					
	1	**2**	**3**	4	5	...
Standard	17,8%	**23,4%**	**26,4%**	29,1%	31,2%	...
Alternative	16,3%	**22,2%**	**24,9%**	27,5%	29,4%	...

Standard coding is better than the alternative one in regard to variance explained.

3-dimensional solution explains just a little bit more variance than the 2-dimensional one, therefore the latter was selected for the visualization of the results.

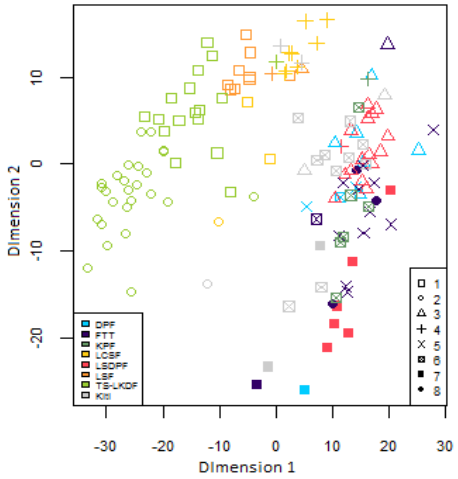

Fig. 1. Manhattan distance, standard coding

The results of the two best clusterings are rendered in the Figs. 1 and 2. Different colors show different factions, and different shape of the points means different clusters.

It is seen in the plots that the location of position and opposition clearly differ. Also, it is interesting that although the unity of both position and opposition is unarguable, it shows differently. Factions in the position show more inner solidarity while keeping a certain distance between them, and no faction in the opposition shows a unique posture comparing to the others, and their inner solidarity is poorer in comparison to position.

Alternative coding of votes "pushes" MPs closer to the center of position/ opposition (this is partially influenced by a smaller amount of different coding values). However, the interpretation of the solution is essentially the same as the one generated using standard coding.

According to the results of clustering we can visualize the predicted assignment of MPs to factions, using Table 5. In this case each cluster would be assigned to the factions, therefore we must take external validation indexes into account. The best values of the external quality evaluation criteria were produced using Manhattan distance and standard coding (Table 9). The prediction are depicted in Fig. 3.

Percentage of correctly assigned to factions MPs is 68,35% (*Purity*, Table 4). While predicting the assignment smaller factions occasionally are overshadowed by larger and their members are assigned to larger factions. The best example would be cluster no. 1 containing 9 of 11 members of LSF, but despite that TS-LKDF is assigned to the aforementioned cluster due to the fact that it contains 16 members of TS-LKDF.

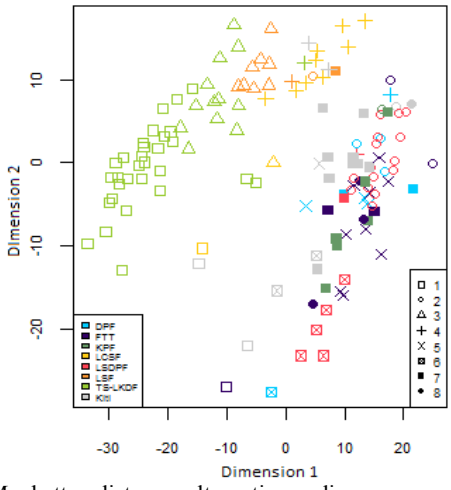

Fig. 2. Manhattan distance, alternative coding

Table 9. Prediction of MPs assignment to factions (Manhattan distance, standard coding)

Cluster	Objects in cluster	Faction assigned	No. of MPs assigned correctly	Actual no. of MPs in faction
1	27	TS-LKDF	16	43
2	28	TS-LKDF	26	43
3	26	LSDPF	15	22
4	14	LCSF	8	11
5	13	FTT	11	17
6	18	Other	10	18
7	10	LSDPF	6	22
8	3	FTT	3	17

5. Results and conclusions

1. **Clustering is a suitable method for parliamentary voting analysis.** The prediction of MPs assignment to factions according to their respective cluster was performed with 68% accuracy. It is a good result, especially when in consideration of the lack of clear and unique arrangement of the opposition factions.
2. **Multidimensional scaling is a useful set of techniques for visualization of parliamentary voting data.** The differences between position and opposition are displayed nearly perfectly. The (un)solidarity of factions can be seen quite well too.
3. Manhattan distance seems to be the most suitable distance with *k*-means method. Alternative votes coding produced clustering of higher quality, however, assuming each cluster should be assigned

to a certain faction, it is the standard coding that appeared more suitable.

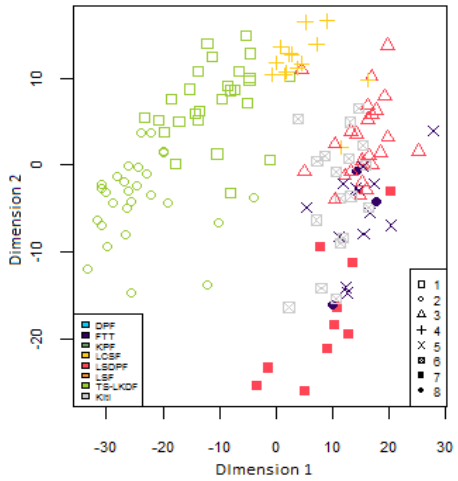

Fig. 3. Prediction of MPs assignment to factions

In accordance with results, several important political conclusions can be made as well.

1. MPs vote quite solidly with their colleagues in the same factions (differences between position and opposition are especially visible).
2. Unity of position and opposition is represented in a different way. Each of the 3 factions in position has a solid location that is a little bit different from the others. Members of opposition factions incline to vote against position, but with no clear location among themselves.
3. It is more difficult to keep unity among the members of a large faction. It is clearly seen in the votings of TS-LKDF and LSDPF (two largest factions).
4. No clear groups, whose members would vote consistently differently from their faction, were detected.

Future research:

1. **Different codings and dissimilarity measures.** Two different codings were examined in depth, but choosing the most suitable coding is trivial and other codings will be analyzed. The same goes with dissimilarity measures – only a small part of them was discussed in this paper.
2. **Different meaning/weight of the votings.** Some of the votings in Seimas are much more important than the others. Hence the results of the research would be much more accurate and objective assigning different importance to every voting.
3. **Software development.** An R package with the options to automate the calculations and visualization of the results is being developed at the time of writing. The goal of the aforementioned package is to quicken the analysis of parliamentary voting significantly.

Research is funded by ESFA (VP1-3.1-ŠMM-10-V-02-025).

6. References

1. Global Politician: http://www.globalpolitician.com/default.asp?26729-democracy-voting-elections.
2. Hix S., Noury A., Roland G. *Dimensions of Politics in the European Parliament.* American Jrnl. of Political Science, 2006. 50(2): 494-520.
3. Hix, S. and Noury, A. *Government-Opposition or Left-Right? The Institutional Determinants of Voting in Fourteen Parliaments.* Working paper, v. 1, 2008.
4. Poole, K. T. *Spatial Models of Parliamentary Voting*. Cambridge University Press. 2005.
5. Morkevičius V. *Neideologinis Seimas? Statistinė svarbių 2004-2008 m. kadencijos Lietuvos Seimo balsavimų analizė. Partinės demokratijos pabaiga? Politinis atstovavimas ir ideologijos*. Vilnius: Versus Aureus, 53-87, 2009.
6. Krilavičius T., Morkevičius V. *Mining Social Science Data: a Study of Voting of Members of the Seimas of Lithuania Using Multidimensional Scaling and Homogeneity Analysis*. 2011.
7. Krilavičius T., Morkevičius V. *Ideological positions of Lithuanian MPs and their dynamics in the term 2008-2012: Statistical Analysis of MPs Voting*. Proc. of Conf. of the Lithuanian Political Science Association, Vilnius, 2010.
8. Užupytė R., Morkevičius V., Krilavičius T. *Lietuvos Respublikos Seimo narių balsavimų tyrimas pasitelkiant socialinių tinklų analizę: tinklo konstravimo parametrų įtaka.* Proc. of students scientific practice conf. 2012.
9. Krilavičius T., Žilinskas A. *On Structural Analysis of Parlamentarian Voting Data.* Informatica 19, 3 2008, 377-390.
10. Krilavičius T., Cimmperman P., Žalandauskas T. *Duomenų užteks visiems [Plenty data for everyone]*. Vilnius, 2010.
11. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Austria. 2012
12. Muenchen R. *The Popularity of Data Analysis Software.* http://r4stats.com/articles/popularity/.
13. Kabacoff R. I. *R in Action. Data analysis and graphics with R.* 2011.
14. Čekanavičius V., Murauskas G. *Statistika ir jos taikymai II.* – Vilnius: TEV, 272 psl., 2002.
15. Basu S., Davidson I., Wagstaff K. L. *Constrained Clustering. Advances in Algorithms, Theory and Applications.* 2008.
16. Manning, Ch.D., Raghavan, Schutze H. *Introductions to Information Retrieval*. Camb. Univ. Press. 2008.
17. Rendon E., Abundez I., Arizmendi A., Quiroz E. *Internal versus External cluster validation indexes.* Int. Journal of Computers and Communications. Issue 1, Volume 5 (2011).
18. Trinity University: http://www.trinity.edu/rjensen/352wpvisual/000datavisualization.htm
19. Abdi H. *Metric Multidimensional Scaling (MDS): Analyzing Distance Matrices*. URL: http://www.utd.edu/~herve/Abdi-MDS2007-pretty.pdf
20. Buja A., Swayne D. F., Littman M. L., Dean N., Hofmann H., Chen L. *Data visualization with multidimensional scaling.* 2008.