Tensor Decomposition for Multi-way Time-Varying Data Visualization

Luzia de Menezes Romanetto*, Aline Aparecida de Souza Leão**, Fabio Dias* and Luis Gustavo Nonato* *Member, IEEE*
* ICMC - Universidade de Sao Paulo, São Carlos, Brazil
** UTFPR - Federal Technological University of Paraná, Cornélio Procópio, Brazil
luzia@icmc.usp.br, aasleao@gmail.com, fabio.dias@gmail.com, gnonato@icmc.usp.br

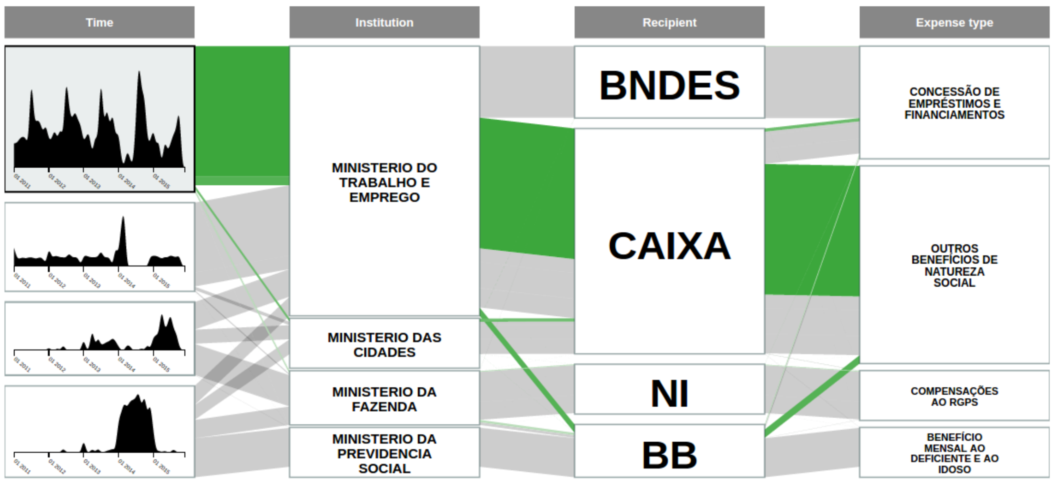


Fig. 1: The proposed visualization interface containing spending data from the Brazilian Federal Government. Our method correctly identifies and displays a relevant temporal pattern mainly related to a common policy of the Labour Ministry to pay workers benefits, that is periodic over years and have a peak in the middle of the year.

***Abstract*—Recent developments in multi-way array decomposition, also known as tensor decomposition methods, have made the analysis and interpretation of patterns in multi-way data a feasible task, fostering a multitude of applications in machine learning and data analysis. However, the use of tensor decomposition to assist visualization applications is still incipient, mainly in problems involving time-varying data. In this work, we present a time-varying multi-way data visualization method that relies on tensor decomposition to reveal hidden temporal patterns as well as the data components that most contribute for those patterns.**

I. Introduction

In the last decade multi-way analysis has gained popularity, mainly due to the theoretical and computational advances in tensor decomposition methods [1]. In fact, multi-way analysis via tensor decomposition has been successfully employed in problems involving audio and video processing, machine learning, and data mining [2]. Nevertheless, the use of tensor decomposition in the context of visualization is still incipient. Existing techniques rely on tensor decomposition mostly for social network visual analysis and glyph-based visualization (see Subsection I-A), thus not exploiting the potential of multi-way representations to handle time-varying data together with the relationship among the tensor modes.

In this work we show how the factors resulting from a non-negative tensor decomposition scheme can be interpreted and built into a parallel sets [3] visual metaphor to enable a simple and intuitive visual analysis of time-varying multi-way data. We also employ interactive resources to illustrate how the different modes of the tensor relate to each other, allowing the visual analysis of the patterns hidden in the data. Moreover, an optimization procedure is proposed to reduce edge crossing, typically found in parallel sets visual metaphor, thus reducing visual clutter. Results from a real data case study show the effectiveness of our approach in revealing gist information from complex data sets.

In summary, the main contributions of this work are:

- A novel time varying data visual analysis methodology that relies on tensor decomposition, pattern recognition, and parallel sets visual metaphor.
- An optimization procedure to reduce edge crossing in parallel sets, making the visual metaphor more scalable for multiple modes.
- A visualization system to assist users in the exploration of time-varying data represented as multi-way arrays, which builds upon a new interpretation of tensor decomposition components to reveal patterns hidden in the data.

A. Related work

In this section we discuss techniques that, at some stage of the visualization process, demand a multi-way representation and/or tensor decomposition to produce visualizations. Special attention is given to visualization techniques where temporal information is considered as one of the modes in a multi-way representation. A comprehensive discussion about time-varying data visualization is beyond the scope of the paper and we refer interested readers to the surveys of Aigner et al. [4] and Kehrer and Hauser [5].

Mosaic Plots and its variants [6], [7] figure among the main options for visualizing multi-way data. However, those plots do not scale properly beyond five or six modes and they also present limitations when dealing with time-varying multi-way data. As an alternative, Cox and Hackborn [8] propose a method that uses one view for each dimension of a multi-way table, linking distinct views such that user actions in one view are reflected in the others.

The combination of multi-way representation and tensor decomposition has been exploited in the context of visual analysis of social networks. Oliveira and Gama [9] propose the use of 3-way arrays to visualize students friendship networks. Tensors are built using statistics of each "actor" in fixed time slices, setting time as the third dimension. The visualization is accomplished by properly projecting tensor modes onto the subspace generated by the two most representative components of each factor resulting from the the tensor decomposition. However, the two-dimensional scatter plot does not favor the visual analysis of the interplay between different tensor modes nor the identification of temporal patterns.

Tensor decomposition has also been employed in the context of simulations. Ballester-Ripoll [10] use the tensor train representation to visualize the parameter space for varied automotive simulations, applying this tensor format to enable an exploratory visualization of the parameter space in complex simulations.

In general, the methods above only visualize one or two modes at once or visualize different modes separately, fragmenting the visual analysis and thus demanding larger cognitive effort to understand patterns and phenomena hidden in the data. In contrast, our approach makes use of parallel sets to visually connect the different modes resulting from a tensor decomposition, giving the user a full picture of the latent information in the data.

II. Nonnegative Tucker Decomposition and Its Interpretation

An *n-tensor* $\mathcal{T}$ is a multi-way array with elements indexed by an n-tuple of indices, that is, each entry in $\mathcal{T}$ is represented as $t_{i_1,i_2,\ldots,i_n}$. Each index i_k corresponds to a *mode* of the tensor with range in the set $\{1, 2, \ldots, I_k\}$, where I_k is the *dimension* of mode k. Scalars, vectors, and matrices are particular cases of 0-tensors, 1-tensors, and 2-tensors, respectively.

Similarly to matrices, tensors can be decomposed as simpler tensors (for example, rank one tensors whose precise definition is beyond our scope, see [2] for details), each one containing relevant information about the original tensor. Finding such simpler tensors is the end goal of the so-called tensor decomposition methods. There are several classes of tensor decomposition methods, being Canonical Polyadic decompositions [11], Tucker decompositions [2], and Tensor-train decomposition [12] the most important ones. Tucker decompositions figure among the most flexible alternatives, mainly in terms of constraints that can be imposed during the decomposition process. Therefore, we have chosen the Tucker decomposition as basis for the present work.

Tucker decompositions take as input an n-tensor $\mathcal{T} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_n}$ and output a *core tensor* $\mathcal{G} \in \mathbb{R}^{R_1 \times R_2 \times \cdots \times R_n}$ with the same number of modes as $\mathcal{T}$ and a set of n *factor matrices* $\mathbf{A}^{(k)} \in \mathbb{R}^{I_k \times R_k}$, one for each mode of $\mathcal{T}$. By properly combining the core tensor $\mathcal{G}$ and the factor matrices, one can reconstruct or approximate the original tensor $\mathcal{T}$. The representation provided by a Tucker decomposition is given by the following expression:

$$\mathcal{T} \approxeq \sum_{i_1=1}^{R_1} \sum_{i_2=1}^{R_2} \cdots \sum_{i_n=1}^{R_n} g_{i_1,i_2,\ldots,i_n} \left(\mathbf{a}_{i_1}^{(1)} \circ \mathbf{a}_{i_2}^{(2)} \circ \cdots \circ \mathbf{a}_{i_n}^{(n)} \right) \tag{1}$$

where $g_{i_1,i_2,\ldots,i_n}$ is an entry in $\mathcal{G}$, $\mathbf{a}_{i_k}^{(k)}$ the i_k-th column of the factor matrix $\mathbf{A}^{(k)}$, $\mathbf{a}_{i_k}^{(k)} \circ \mathbf{a}_{i_l}^{(l)}$ is the outer product between the column vectors $\mathbf{a}_{i_k}^{(k)}$ and $\mathbf{a}_{i_l}^{(l)}$, and the values of R_i are chosen to achieve a compromise between the extraction of relevant information and computational cost. Figure 2 illustrates how the first term of the summation in Equation (1) is obtained.

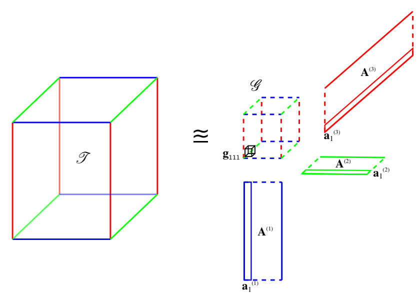


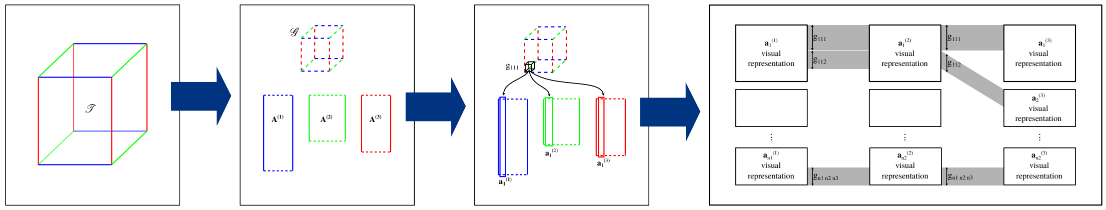
Fig. 2: Illustration of the Tucker decomposition of a 3-tensor $\mathcal{T}$, resulting in a core tensor $\mathcal{G}$ and three factor matrices $\{\mathbf{A}^{(1)}, \mathbf{A}^{(2)}, \mathbf{A}^{(3)}\}$.

In the context of the present work we consider the Nonnegative Tucker Decomposition (NTD) [13], where the values on the core tensor and the factor matrices are enforced to be nonnegative. Non-negativity constraints naturally introduce sparsity and a "part" based representation of the data, making it easier to interpret results [14]. Moreover, the magnitude of the values in the core tensor are directly related to relevance of the factors and their interplay.

A. Tucker interpretation

Equation 1 tells us that the original tensor $\mathcal{T}$ is decomposed as a linear combination of simpler tensors $\mathcal{C}_{i_1,i_2,\ldots,i_n} = \mathbf{a}_{i_1}^{(1)} \circ \mathbf{a}_{i_2}^{(2)} \circ \cdots \circ \mathbf{a}_{i_n}^{(n)}$ given by the outer product of the columns of the factor matrices, scaled by the corresponding element of the core tensor $g_{i_1,i_2,\ldots,i_n}$. Each of those simpler tensors comprises a piece of latent information within the original data tensor.

From this point of view, it is similar to what is given by the SVD matrix decomposition, where we have the singular



Tensorial data Tensor decomposition NTD Relationship between the tensor core and factor matrices Visualization by parallel sets

Fig. 3: Visualization pipeline.

vectors combined and scaled by their singular value. In that decomposition the singular vectors can be seen as pattern that defines a piece of the space where it is. Similarly, in this work we interpret each column $\mathbf{a}_{i_j}^{(j)}$ as composing patterns in the corresponding modes, i. e., in the case of a tensor formed by modes of (time $\times$ user $\times$ item), the columns can be interpret as patterns of time, patterns of users and patterns of items, respectively. At same time, we interpret the value $g_{i_1,i_2,\ldots,i_n}$ as the importance of the information constructed by the combination of the patterns $\{\mathbf{a}_{i_1}^{(1)}, \mathbf{a}_{i_2}^{(2)}, \ldots, \mathbf{a}_{i_n}^{(n)}\}$, i.e., the entries $g_{i_1,i_2,\ldots,i_n}$ of the core tensor reveals which set of patterns are more relevant and more connected.

The interpretation above is the building block of our visualization method, as detailed in section III.

III. Visualization Layout

Following the presented Tucker interpretation we propose a visualization based on a variant of parallel sets [3]. Our approach follows the pipeline illustrated in Fig. 3.

A. Parallel Sets representing tensor decomposition

Parallel sets [3] are specially designed to represent data with categorical dimensions, meaning that in each dimension there are a limited number of different values. Considering a dataset with n categorical attributes, each attribute is represented by a pillar, which are organized side by side similarly to parallel coordinates. The pillars are segmented in bins, one for each possible value the corresponding attribute can assume. The bins between adjacent pillars are connected if there are instances with attributes equal to the corresponding bins. The larger the number of instances sharing the corresponding attributes, the thicker the line strip connecting the bins.

In this work we use the foundations of parallel set as visual metaphor, letting the tensor modes to play the role of pillars and the discovered patterns (columns from the factor matrices) as the categories. More specifically, each factor matrix gives rise to a pillar in the parallel sets metaphor. The bins associated to each pillar correspond to the columns of the factor matrix. For instance, the layout depicted in the right of Fig. 3 shows the decomposition of a 3-tensor. In each bin a corresponding pattern is visually expressed by a appropriated metaphor. Since in this work we focus on temporal pattern the leftmost pillar always corresponds to a temporal mode of the tensor, and will be expressed as a time series. The middle and rightmost pillars correspond to the other two modes of the tensor. The number of bins in each column is dictated by the dimension R_i chosen in the decomposition. The values of the core tensor $g_{i_1,i_2,\ldots,i_n}$ are expressed by the line strips, i. e., the larger the value of $g_{i_1,i_2,\ldots,i_n}$ is the wider the related line strip. These components together are capable of expressing all values given by a tensor decomposition.

To help users better understand the data, in our visualization we implement a number of visual and interactive resources, including color codes, highlight and selection. Users can also select some patterns of interest by clicking in particular bins or line strips, as depicted in Fig. 1.

B. Crossing minimization problem in weighted multi-layer graph

The growth in the number of modes and patterns increases the number of elements to be represented in the layout, so the visualization can become cluttered due to the amount of crossing between line strips. To avoid this problem we implemented two strategies: relaxation of the strips input/output order, and crossing minimization between adjacent pillars.

Based on [3], we allow the input and output position of the line strip in each bin to be different, that is, the vertical position a line strip reaches a particular bin from the left does not need to be the same it leaves the bin on the right. Such relaxation reduces the number of edge crossing.

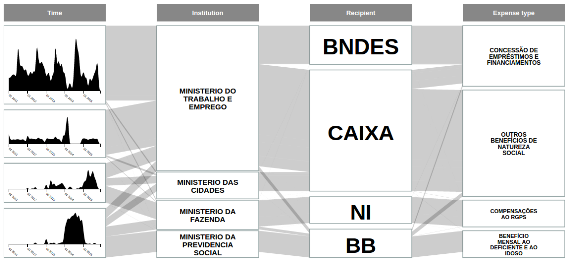
To solve the crossing minimization problem in weighted multi-layer graph we used the integer programming model proposed by Junger [15], but considering weighted edges. As a result, the problem is to minimize the crossing edges giving higher priority to edges with larger weight. Given the space limitation, it will not be explained here.

IV. Results

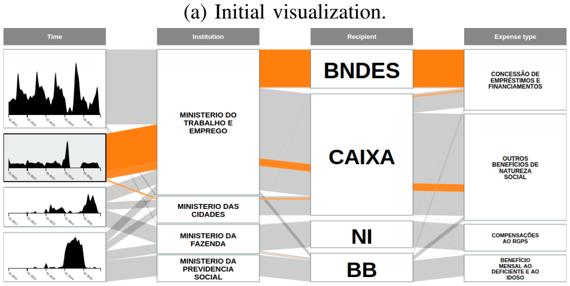
The effectiveness of the proposed methodology is shown in a real application scenario, a case study involving large amount of Brazilian government data.

A. Visualizing Brazilian Government Spending

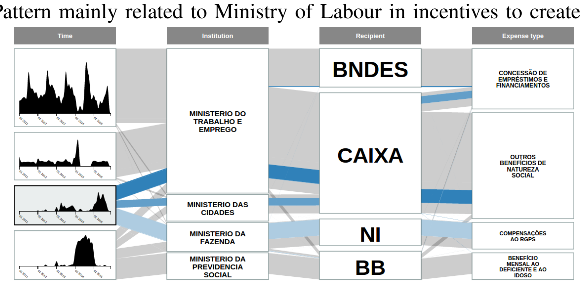
For this case study, we considered expense information from the transparency initiative of the Brazilian government [16], containing 60 millions transactions, from January 2011 to October 2015, for a total value just over R$ 1 trillion. Four modes were considered: temporal, the ministry which authorized it, the recipient type, and the type of the expense. Since most of the recipient types were banks, we identified the three most common banks directly: the Bank of Brazil (BB), the Federal Savings Bank (CAIXA), and the National Bank for Economic and Social Development (BNDES). Additionally, for several reasons, the recipient of the expenses are not always identified, resulting in a new category Not Informed (NI). Figure 4a illustrates our visualization of this dataset, considering a resulting core tensor of dimensions $(4, 4, 4, 4)$. The dimensions have been obtained by the DIFFIT method [17].



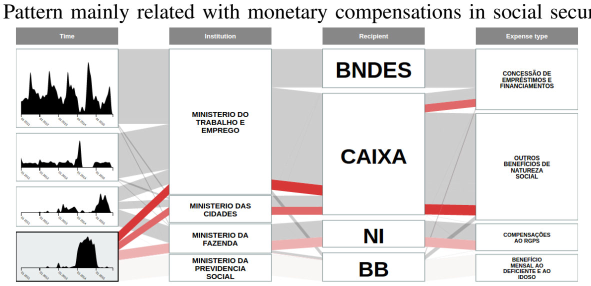
(a) Initial visualization.



(b) Pattern mainly related to Ministry of Labour in incentives to create jobs.



(c) Pattern mainly related with monetary compensations in social security.



(d) Pattern mainly related to a special policy that was active only in 2014.

Fig. 4: Visualization of spending patterns in the initial visualization and highlighting relation between temporal patterns and other modes.

Starting from the temporal patterns, we identify four distinct spending profiles, where it's easy to see common temporal patterns as periodic, seasonal, and peaks (first pillar in Fig. 4). Since we are working with governmental data, each of this patterns are associated with political events or known trends. We have investigated their origins based on media reports, which clearly identify the underlying events for each pattern.

In each of the temporal patterns we can observe useful insights based on the activity magnitude:

- The first pattern shown in Figure 1, having mainly the Labour Ministry (Ministério do Trabalho e Emprego) as origin and CAIXA as recipient, is related to worker social benefits, such as unemployment and health insurance. The periodic peaks starting in July are related to the PIS program of social benefits, payed between July/August to October/November between 2011 to 2014. In 2015 the payment calendar was modified to spread the benefit between July 2015 to March 2016, which can be visualized in the same temporal pattern [18].
- The second pattern shown in Figure 4b, having mainly the Labour Ministry as origin and BNDES as recipient, is related to incentives for job creation via loans to private companies. In 2013, the ministry had a large debt with the bank, which was paid at the beginning of 2014, which is observed as a peak in the pattern. It was reported that there was no other payments during this year, which is also observed in the pattern.
- The third pattern shown in Figure 4c, presents some mixing with the previous and next pattern, but its main characteristics are payments occurring during 2013 and 2015 from the Treasury (Ministério da Fazenda) to expenses to a social security fund. This is a special group for accounting expenses such as rate reductions and exemptions.
- The last group of expenses is depicted on Figure 4d. It is mainly comprised of benefits for the elderly and disabled, from the Ministry of Social Security (Ministério da Previdência Social). Further, this expense occurred mostly in 2014. Indeed, this expense is provenient from a different interpretation of the Brazilian Constitution, implemented in 2014, that was later deemed unconstitutional and interrupted in 2015.

Additionally, we can view by color coding that the Ministry of Cities (Ministério das Cidades) has a connection to CAIXA for loan concession spending for all figures. This correlates to social habitation programs and subsides offered by the federal government that do not have a strong temporal characteristic. As such, the factorization "mixed" this type of spending into each temporal pattern. This can be viewed by selecting the corresponding bin in the interactive visualization, but not included due to space limitations.

Finally, it's worth highlighting that the bin size and time series amplitudes do not map to a monetary value, but rather to a relative importance measure.

V. Conclusion

In this work we have proposed a novel visual analytic methodology for analyzing multi-way time-varying data that combines tensor theory, pattern recognition, and parallel sets visual metaphor in an interactive visualization. Also, it transforms a complex structure such as a tensor into a visual representation that includes all extracted information from a tensorial decomposition technique, summarizing the information in an accessible form, effective in revealing important events given by variation in the data. The usefulness of the proposed methodology was attested in a real data case study.

In our visualization it is possible to easily identify the connection between the temporal pattern and its behavior in other modes of the data, as well as the importance of each pattern relative to the whole dataset as expressed by its magnitude. Due to the non negativity restriction imposed, the more relevant patterns are detected in the decomposition and presented in the visualization, which can give insights to guide the knowledge extraction by presenting only the most important information.

ACKNOWLEDGMENT

The authors acknowledge financial support from Coordenação de Aperfeiçoamento de Pessoal de Nıvel Superior (CAPES), Grants 2014/12815-1, 2016/04391-2, 2011/22749-8 São Paulo Research Foundation (FAPESP) and the National Council of Technological and Scientific Development (CNPq) for the financial support of this work. The views expressed are those of the authors and do not reflect the official policy or position of the São Paulo Research Foundation.

REFERENCES

[1] A. Cichocki, D. Mandic, L. De Lathauwer, G. Zhou, Q. Zhao, C. Caiafa, and H. A. Phan, "Tensor decompositions for signal processing applications: From two-way to multiway component analysis," *Signal Processing Magazine, IEEE*, vol. 32, no. 2, pp. 145–163, 2015.

[2] T. G. Kolda and B. W. Bader, "Tensor decompositions and applications," *SIAM review*, vol. 51, no. 3, pp. 455–500, 2009.

[3] R. Kosara, F. Bendix, and H. Hauser, "Parallel sets: Interactive exploration and visual analysis of categorical data," *Visualization and Computer Graphics, IEEE Transactions on*, vol. 12, no. 4, pp. 558–568, 2006.

[4] W. Aigner, S. Miksch, W. Müller, H. Schumann, and C. Tominski, "Visualizing time-oriented data - A systematic view," *Computers & Graphics*, vol. 31, no. 3, pp. 401–409, 2007.

[5] J. Kehrer and H. Hauser, "Visualization and visual analysis of multifaceted scientific data: a survey." *IEEE Trans. Vis. Comput. Graphics*, vol. 19, no. 3, pp. 495–513, 2013.

[6] M. Friendly, "A brief history of the mosaic display," *Journal of Computational and Graphical Statistics*, vol. 11, no. 1, pp. 89–107, 2002.

[7] K. Hornik, A. Zeileis, and D. Meyer, "The strucplot framework: visualizing multi-way contingency tables with vcd," *Journal of Statistical Software*, vol. 17, no. 3, pp. 1–48, 2006.

[8] K. C. Cox and D. Hackborn, "Interactive visualization of multiway tables," in *infovis*. IEEE, 1996, p. 68.

[9] M. Oliveira and J. Gama, "Visualization of evolving social networks using actor-level and community-level trajectories," *Expert Systems*, vol. 30, no. 4, pp. 306–319, 2013.

[10] R. Ballester-Ripoll, E. Paredes, and R. Pajarola, "A surrogate visualization model using the tensor train format," in *SIGGRAPH ASIA 2016 Symposium on Visualization*. ACM, 2016, p. 13.

[11] R. A. Harshman, "Foundations of the parafac procedure: models and conditions for an" explanatory" multimodal factor analysis," 1970.

[12] I. V. Oseledets, "Tensor-train decomposition," *SIAM Journal on Scientific Computing*, vol. 33, no. 5, pp. 2295–2317, 2011.

[13] A. Cichocki, R. Zdunek, A. H. Phan, and S.-i. Amari, *Nonnegative matrix and tensor factorizations: applications to exploratory multi-way data analysis and blind source separation*. John Wiley & Sons, 2009.

[14] D. D. Lee and H. S. Seung, "Learning the parts of objects by non-negative matrix factorization," *Nature*, vol. 401, no. 6755, pp. 788–791, 1999.

[15] M. Jünger, E. K. Lee, P. Mutzel, and T. Odenthal, "A polyhedral approach to the multi-layer crossing minimization problem," in *Proceedings of the Symposium on Graph Drawing, Lecture Notes in Computer Science 1353*. Springer, 1997, pp. 13 – 24.

[16] "Portal da Transparência — Governo Federal," http://portaldatransparencia.gov.br/, 2016.

[17] M. E. Timmerman and H. A. Kiers, "Three-mode principal components analysis: Choosing the numbers of components and sensitivity to local optima," *British journal of mathematical and statistical psychology*, vol. 53, no. 1, pp. 1–16, 2000.

[18] "PIS - Programa Integração Social — Caixa," https://goo.gl/dn9qwa, 2017, [Online; accessed 04-March-2017].