

RESEARCH ARTICLE

Time-Varying Transition Probability Matrix Estimation and Its Application to Brand Share Analysis

Tomoaki Chiba¹, Hideitsu Hino^{2*}, Shotaro Akaho³, Noboru Murata¹

1 Department of Electrical Engineering and Bioscience, Waseda University, Shinjuku, Tokyo, Japan, **2** Department of Computer Science, University of Tsukuba, Tsukuba, Ibaraki, Japan, **3** Mathematical Neuroinformatics Group, National Institute of Advanced Industrial Science and Technology, Tsukuba, Ibaraki, Japan

* hinohide@cs.tsukuba.ac.jp



OPEN ACCESS

Citation: Chiba T, Hino H, Akaho S, Murata N (2017) Time-Varying Transition Probability Matrix Estimation and Its Application to Brand Share Analysis. PLoS ONE 12(1): e0169981. doi:10.1371/journal.pone.0169981

Editor: Yongtang Shi, Nankai University, CHINA

Received: June 1, 2016

Accepted: December 26, 2016

Published: January 11, 2017

Copyright: © 2017 Chiba et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper and its Supporting Information files. Data are also available from author's web page: http://www.cs.tsukuba.ac.jp/~hinohide/code/brand_analysis.zip.

Funding: This work was supported by JSPS KAKENHI Grant Numbers 25120009, 25120011, and 16K16108. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

Abstract

In a product market or stock market, different products or stocks compete for the same consumers or purchasers. We propose a method to estimate the time-varying transition matrix of the product share using a multivariate time series of the product share. The method is based on the assumption that each of the observed time series of shares is a stationary distribution of the underlying Markov processes characterized by transition probability matrices. We estimate transition probability matrices for every observation under natural assumptions. We demonstrate, on a real-world dataset of the share of automobiles, that the proposed method can find intrinsic transition of shares. The resulting transition matrices reveal interesting phenomena, for example, the change in flows between TOYOTA group and GM group for the fiscal year where TOYOTA group's sales beat GM's sales, which is a reasonable scenario.

Introduction

Multivariate time series recording of actual phenomenon may have dynamics based on an intrinsic variable structure. In particular, we consider the transition probabilities among products such as beer, automobiles, and newspapers. The transition matrix describes the probability of a change from one state to another state. In concrete terms, the transition matrix characterizes the shifts in consumers' preferences towards different products in terms of probability. In other words, we assume that the (i, j) -element of the transition matrix is the probability that a consumer who previously bought the i -th product now purchases the j -th product instead.

Information on changes in consumers' product purchases is necessary for understanding competition between products in the market. The intrinsic structure of transition matrices, however, cannot be directly observed because individuals' purchasing data are typically difficult to obtain. Individuals' data are also difficult to handle because of privacy issues. Owing to these issues, marketing data are often limited to product sales amounts such as point-of-sales

data, aggregated by removing personally identifiable information. Therefore, a method to estimate and analyze the structure based on observations is indispensable for understanding hidden dynamics such as product share.

The method proposed in this paper estimates the transition matrices of customers switching between products by using only aggregated sales share data. Although it is impossible to uniquely determine the transition matrix from sales share data only, by making some natural assumptions on the transition matrix, we propose a method to estimate all of the transition matrices for every observation.

For the purposes of the mathematical formulation, we introduce the terminology of graph theory [1]. We identify products as nodes, and the relations between products are expressed by the edges between nodes. Consumers are supposed to move between nodes under the condition that the total number of consumers in the entire graph before and after the movement is constant. This consumer behavior is thus modeled as a finite state Markov chain with the constraint that the total number of consumers is fixed before and after the state transition [2]. The finite state Markov model is popular and widely used, for example, for modeling implied volatility in financial engineering [3], the distribution of individuals in community ecology [4], the distribution of the urban scale in demography [5], and the importance of websites, which is known as the PageRank model [6–8].

Suppose we observe non-negative multivariate time series such as the sales amount of different products. Let us normalize the observed non-negative multivariate time series data so that the sum of all variables is equal to one at each observation. Then, the normalized multivariate π_t at each observation time t is assumed to be the stationary distribution of the consumers on nodes. The stationary distribution $\pi_t \in \mathbb{R}^n$ is a probability vector, which characterizes a transition matrix G_t at time t as $\pi_t^\top G_t = \pi_t^\top$. Then, we consider the problem of estimating the transition matrices corresponding to the observed stationary distributions.

In real problems, it is natural to assume that the intrinsic structure or consumers' preferences vary over time. It is also naturally assumed that the transition probability varies at a slow pace. Additionally, we assume that an observation at time t is the stationary distribution of a transition probability matrix G_t at time t . Hence, we assume that the transition speed of the Markov chain induced by G_t is sufficiently fast.

The analysis of the relationship between variables in multivariate time series is of practical importance in many scientific fields and is also used in social and business data analysis [9–12]. Typically, to analyze the intrinsic structure between variables, it is useful to express the system as a graph composed of variables with nodes and edges representing their relationship. To estimate the intrinsic structure in the variables, causality analysis is a standard approach [13, 14]. This approach is actively studied particularly in econometrics [15]. These methods have been successfully applied to many problems, although they rely on statistical tests for every combination of variables and are computationally demanding. These methods also require a large number of observations to construct a statistical model and to perform a statistical test based on the model. With an increasing number of variables, many computationally efficient methods for estimating the covariance structure have been developed. One representative method is the graphical lasso [16], which is based on a sparse regularization and optimization algorithm. By using the graphical lasso and its variants, methods of analysing the time-varying graph structure are proposed and applied to change point detection [17–19], for example. These methods assume that the observed data are realizations of multivariate Gaussian distributions and suffer from low estimation accuracy for non-Gaussian behaviors in real-world problems. These methods are also unable to identify the asymmetric relationships between the variables. To overcome these problems, we develop a method of estimating the transition

matrix without making an assumption about the distribution of the underlying covariates. The main contributions of this study can be summarized as follows:

- We propose a method for estimating consumer transitions between products at any moment by using sales share data only. By using our method, we can avoid the detailed recording of consumer transitions, which is high in cost or even impossible in reality.
- We apply our method to analyze consumer transitions for automobiles and provide a way in which to infer the change in consumers' preferences towards different manufacturers. The result is reasonable and explains actual social/market events.

Materials and Methods

Material

Data to be analyzed are automobile sales data from the year 2007 to the year 2015 from various countries in quarterly units for each manufacturer. In the automotive industry, the positioning or branding of each manufacturer would gradually change, and the market share is assumed to be in a stationary state. Namely, we consider two different timescales. Within the quarterly unit, the transition of the consumers' preferences are sufficiently rapid, that is, the transition of the consumers' preferences is assumed to be in the stationary state. On the other hand, for a longer timescale, the change in the consumers' preferences is assumed to be slow, and the underlying graph structures would gradually change.

For the sake of simplicity in analysis and visualization, among all manufacturers, the top 14 sellers (BMW Group, Chrysler Group, Daimler Group, FCA (Fiat Chrysler Automobiles), Ford Group, GM Group, PSA (Peugeot Société Anonyme), Renault-Nissan, VW Group, Suzuki, Toyota Group, Honda, Mazda, and Hyundai-Kia Group) are used singly, and other manufacturers are grouped and named "Others." Then, the row sales data are transformed to the form of "sales share," namely, the amount of sales is normalized to the ratio, which is regarded as a stationary distribution at a certain quarterly unit. The share data used in this paper are shown in Fig 1 as a stacked graph.

Model Formulation

In this section, with some assumptions for changes in customers' preferences, we model the relationship between the time series of the transition probability matrix and the share of products. Suppose there are n different products, and we observe a series of ratios or shares $\{\pi_t\}_{t=1}^T$, $\pi_t \in \mathbb{R}^n$ of those products, namely, π_t represents the share of n products at time t . The ratio π_t satisfies conditions

$$\sum_{i=1}^n (\pi_t)_i = 1, (\pi_t)_i \geq 0, \forall i \in \{1, \dots, n\} \tag{1}$$

where $(\cdot)_i$ is the i -th element of a vector. From the Perron-Frobenius theorem [20], π_t is considered an eigenvector of a stochastic matrix, namely, π_t is an eigenvector of a transition probability matrix. The transition probability matrix of consumers' preferences on manufacturers at time t is denoted by $G_t \in \mathbb{R}^{n \times n}$, where the (i, j) element of the matrix G_t , which is denoted by $(G_t)_{ij}$, is the transition probability from the i -th product to the j -th product in a time interval $(t - 1, t]$. We assume that each element of the matrix G_t for all $t \in \{1, \dots, T\}$ is strictly positive to account for the probability of random choice by consumers:

$$0 < (G_t)_{ij} \leq 1, \forall i, j \in \{1, \dots, n\}. \tag{2}$$

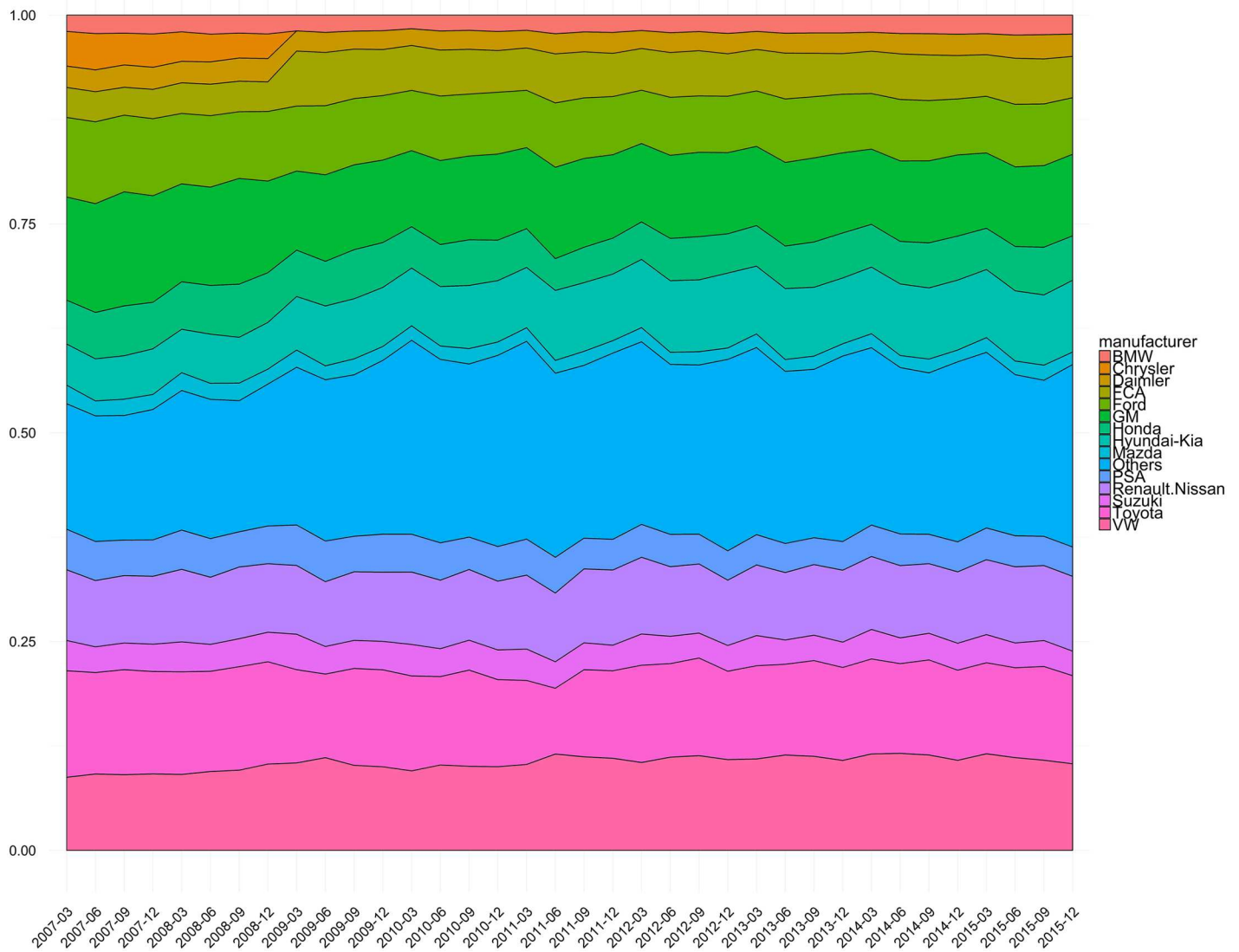


Fig 1. Quarterly sales share of manufacturers from the year 2007 to the year 2015.

doi:10.1371/journal.pone.0169981.g001

Now our problem is estimating a set of transition matrices $\{G_t\}_{t=1}^T$ from observed ratios $\{\pi_t\}_{t=1}^T$, however, this estimation problem is typically indeterminate because the degree of freedom of $\{G_t\}_{t=1}^T$ is greater than that of observations $\{\pi_t\}_{t=1}^T$. Therefore, we need additional constraints on this problem. We impose the following two assumptions.

The first assumption is that the observed ratio π_t at time t is a realization of the stationary distribution of a Markov process represented by G_t . Homogeneous Markov chain modeling based on the stationarity at the observation interval has a long history in marketing research [21]. From research on PageRank, it is also acknowledged that the convergence of distribution π_b by the action of transition matrix G_b to the stationary distribution is very fast [22]. Fig 2 shows a simple example of the transition of distribution π by the consequent actions of transition matrix G . The number of nodes is set to 15 and we show the first 12 time steps of the sequence, namely $\{\pi, G\pi, G^2\pi, \dots, G^{11}\pi\}$. In this figure, the vertical axis shows the node

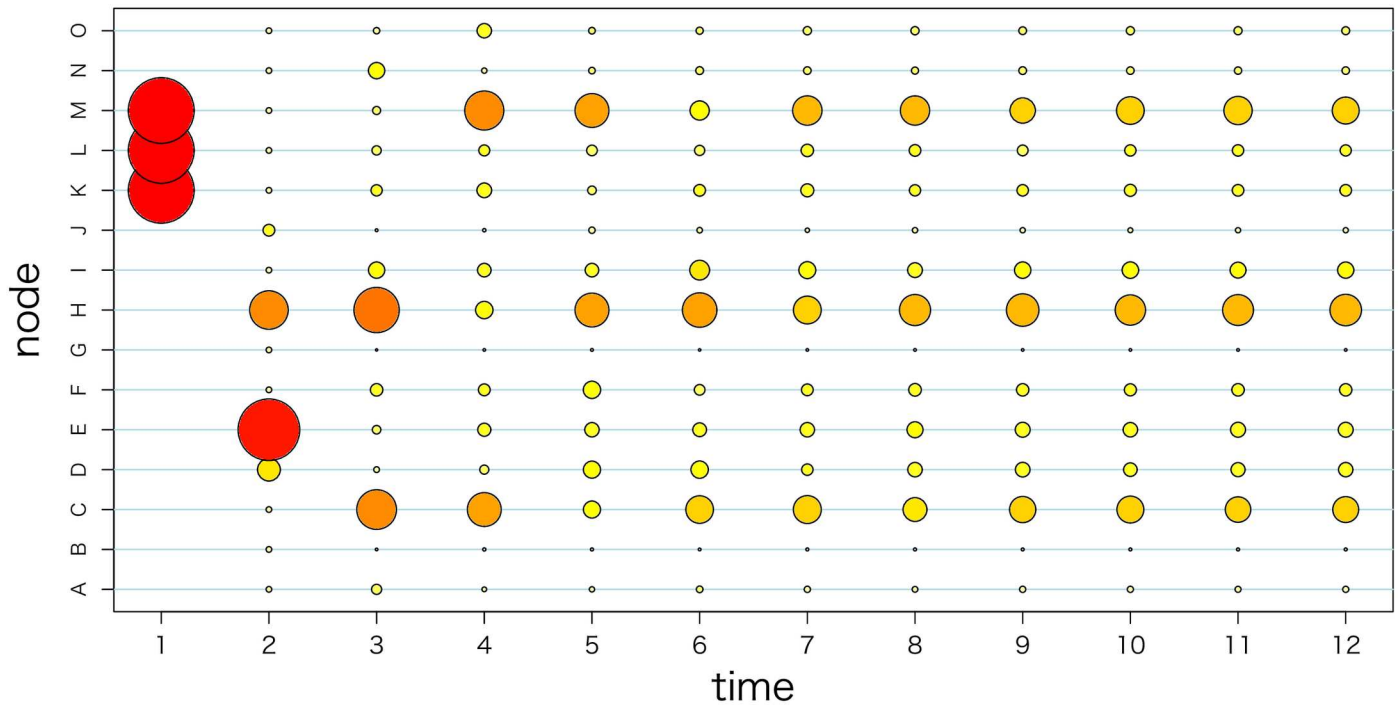


Fig 2. Example of the transition of distribution π through the actions of the transition matrix.

doi:10.1371/journal.pone.0169981.g002

number, which represents the elements of the vector $\pi \in \mathbb{R}^{15}$. The sizes of the circles indicate the relative values of the elements. This figure shows almost no change in the distribution after time step $t = 10$. Because the Markov chain generated by transition matrix G is aperiodic, the stationary distribution is unique regardless of the initial value. In our model, we consider that the quarterly unit is sufficient for the observation to converge to the stationary distribution.

The second assumption is imposed on the time-varying property of transition matrix G_t . Changes in consumers' preferences towards each manufacturer are infrequent, and thus each element of G_t is similar to that of the previous transition matrix G_{t-1} .

We summarize the assumptions on our model below. We treat a sequence of sales share, where each share is quarterly aggregated. We assume that transition matrix G_t in the t -th term is unchanged and that the observed sales share for the term well approximates the stationary distribution of the transition matrix. In other words, the timescale of the transition by the matrix is sufficiently small that the observed share at the end of the term have been converged to the stationary distribution. We note that the *transition* in the case of automobile share is not necessarily the actual purchase of cars by users because it is unusual to buy cars every three months. The transition here indicates the change in *preference* for brands by users, which affects users' buying behavior.

It is also assumed that transition matrix G_t gradually changes and that the stationary distribution of the matrix also changes when observed at different terms. The difference between the consequent transition matrices is thus assumed to be small.

These assumptions about the model of the observed data and transition mechanism are mathematically embodied in the following section, and the problem of estimating the transition matrices is formulated as a simple linear programming.

Optimization Problem for Estimating Transition Matrices

We derive a method for estimating a series of transition matrices $\{G_t\}_{t=1}^T$ corresponding to a series of observations $\{\pi_t\}_{t=1}^T$ of sales shares. We introduce the objective of optimization for estimating G_t and several constraints, which embody the assumptions stated in the previous section.

Assumption 1 (Small and Sparse Change in the Transition Probabilities) *It is natural to assume that changes in transition probability in consecutive observations are not so large. Specifically, we assume that the difference in each element of two consecutive transition matrices is small in terms of the ℓ_1 -norm and define the objective function to be minimized as follows:*

$$L(\{G_t\}_{t=0}^T) = \sum_{t=1}^T \|G_t - G_{t-1}\|_1, \tag{3}$$

where $\|A\|_1, A \in \mathbb{R}^{n \times n}$ is defined by $\sum_{i,j=1}^n |(A)_{ij}|$.

It is worth noting that the ℓ_1 -norm minimization induces a sparse solution [23, 24], and the objective function (3) is called the *fused lasso* in the literature of sparse regularized regression [25]. We also note that we have to include $G_0 \in \mathbb{R}^{n \times n}$ for the target of estimation owing to this assumption. Since G_0 does not have corresponding observation of share, we estimate G_0 but do not give any interpretation for this extra transition matrix.

Condition 1 (Stationary Distribution at Each Observed Time) *The observed product share vector π_t is assumed to be the stationary distribution with respect to the transition matrix G_t , which is formally expressed by the following condition:*

$$\pi_t^\top G_t = \pi_t^\top, \quad \forall t \in \{1, \dots, T\}. \tag{4}$$

Condition 2 (Constraint to be a Transition Probability Matrix) *By definition, the transition probability matrix is strictly positive. In addition, the following equations must be satisfied for G_t to be a transition matrix:*

$$\begin{cases} G_t \mathbf{e} = \mathbf{e}, \\ (G_t)_{ij} > 0, \end{cases} \quad \forall t \in \{0, 1, \dots, T\} \tag{5}$$

where \mathbf{e} is an n dimensional vector with all ones.

Mathematical Programming. Putting together the introduced objective function and constraints, we obtain the following optimization problem:

$$\underset{\{G_t\}}{\text{minimize}} \sum_{t=1}^T \|G_t - G_{t-1}\|_1 \quad \text{s.t.} \quad \begin{cases} \pi_t^\top G_t = \pi_t^\top, & \forall t \in \{1, \dots, T\} \\ G_t \mathbf{e} = \mathbf{e}, & \forall t \in \{0, 1, \dots, T\} \\ (G_t)_{ij} > 0. & \forall t \in \{0, 1, \dots, T\} \end{cases} \tag{6}$$

Optimization. The optimization problem (6) is an instance of linear programming and is efficiently solved using the simplex method or the interior point method [26]. Namely, by introducing an auxiliary matrix $\Xi_t \in \mathbb{R}^{n \times n}$, we can reformulate the problem (6) as linear programming:

$$\underset{\{G_t\}\{\Xi_t\}}{\text{minimize}} \sum_{t=1}^T \sum_{i,j=1}^n (\Xi_t)_{ij}, \tag{7}$$

$$\text{s.t.} \begin{cases} \pi_t^\top G_t = \pi_t^\top, \\ -(\Xi_t)_{ij} < (G_t - G_{t-1})_{ij} < (\Xi_t)_{ij}, \\ G_t \mathbf{e} = \mathbf{e}, \\ (G_t)_{ij} > 0. \end{cases} \quad \begin{matrix} \forall t \in \{1, \dots, T\} \\ \forall t \in \{0, 1, \dots, T\} \end{matrix} \quad (8)$$

The algorithmic description of the proposed method is shown in Algorithm 1.

Algorithm 1 Algorithm for estimating the transition matrices

- Input:** Non-negative multivariate time series.
- Initialization:** Normalized the observed time series to a sequence of stationary distributions $\{\pi_t\}_{t=1}^T$.
- Estimation:** Solve the linear programming Eq (7).
- Output:** Estimated sequence of transition matrices $\{G_t\}_{t=1}^T$.

Results and Discussion

For the quarterly unit automobile sales data of manufacturers from 2007-1Q to 2015-4Q, we performed the preprocessing explained in the Materials and Methods. We note that 1Q, 2Q, 3Q, and 4Q denote the first, second, third, and fourth quarter in a fiscal year. Then, we applied the proposed method for a series of observed share data $\{\pi_t\}_{t=1}^T$, where $t = 1$ corresponds to “2007-1Q,” and $t = T$ corresponds to “2015-4Q,” to estimate a series of transition probability matrices. The optimization problem (7) is solved by the simplex method using the solver for linear programming GLPK [27]. The observed data is shown in Fig 3, which expresses the same information shown in Fig 1. Fig 1 is popular for showing market information, while Fig 3 is more intuitive in terms of market share transition for a certain time unit. The size and color of circles indicate the proportion of automobile sales for different manufactures. Fig 4 shows the

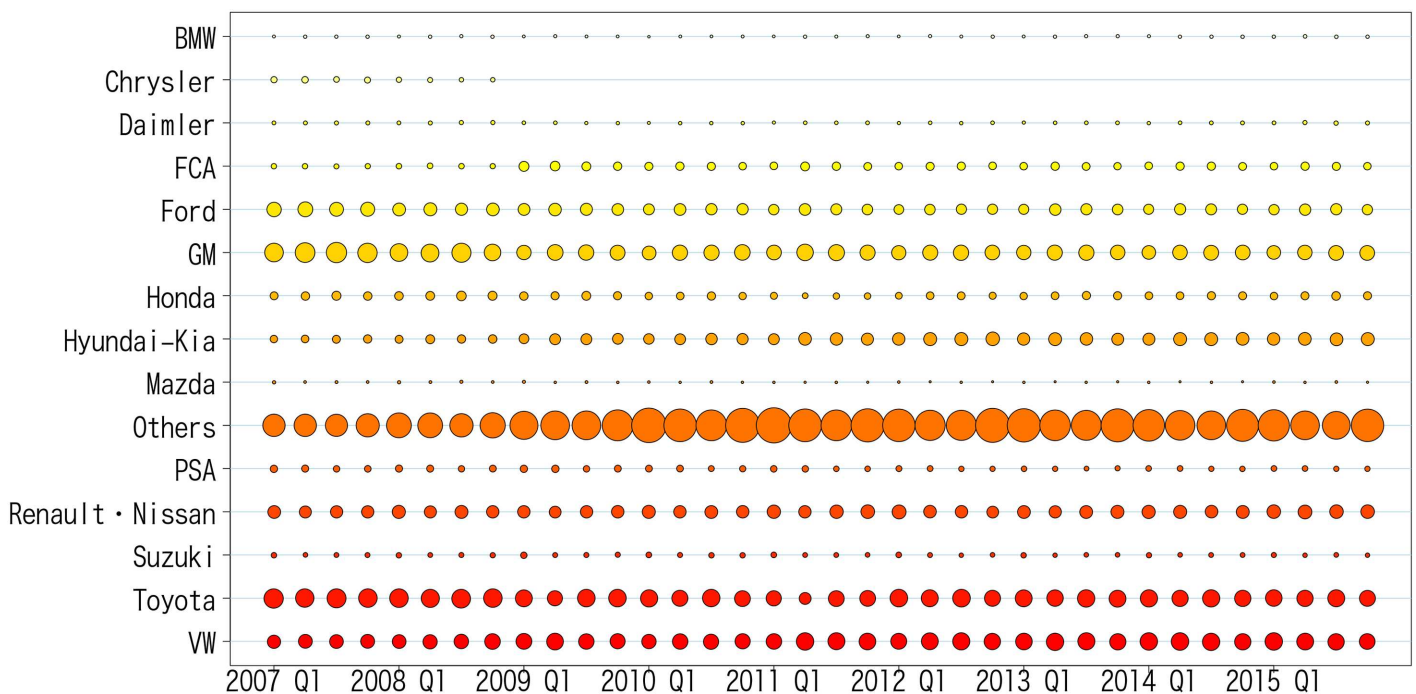


Fig 3. Observed sales share data of manufacturers from the year 2007 to the year 2015.

doi:10.1371/journal.pone.0169981.g003

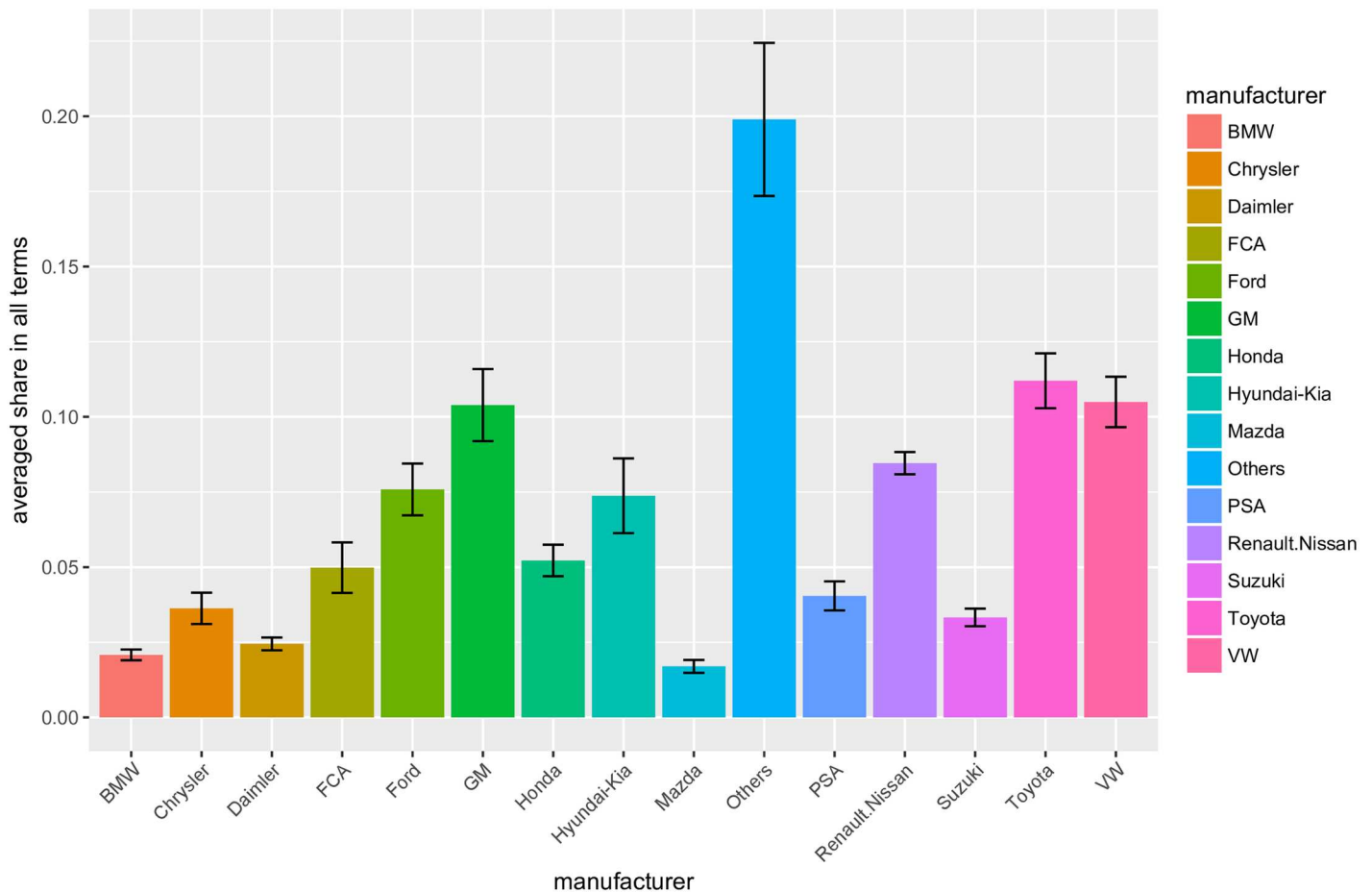


Fig 4. Averages and standard deviations of the sales shares.

doi:10.1371/journal.pone.0169981.g004

averages and standard deviations of the sales shares in all terms. This figure shows that the standard deviations of shares tend to be large for manufacturers with a large market share. Further, the Renault-Nissan group has a relatively high average and small standard deviation, which indicates that this manufacturer maintains a certain market share stably. On the contrary, the Hyundai-Kia group has a relatively low average but a large standard deviation. This fact suggests that this manufacturer is growing rapidly (see also Fig 3). While we can infer the above-mentioned facts and tendencies from the market share data, it is impossible to identify how the consumer transitions from one manufacturer to another with sales share data alone. In the following subsections, we show the graphs representing the estimated transition paths with discussion on and consideration for social events that may explain the estimated results.

Estimated Transition Matrices and Corresponding Market Structure

Transition matrices are constrained to be positive and there are flows of customers from any manufacturer to any other manufacturer. We hereafter set elements of the estimated transition matrices below certain threshold to zero to remove minor edges for visualization purpose. The threshold used to visualize the results in this paper is 0.24, which offers legible results. Fig 5 shows the estimated transition matrix G_1 obtained by solving the linear programming Eq (7), and the corresponding directed graph of automobile manufacturers' sales shares for the first

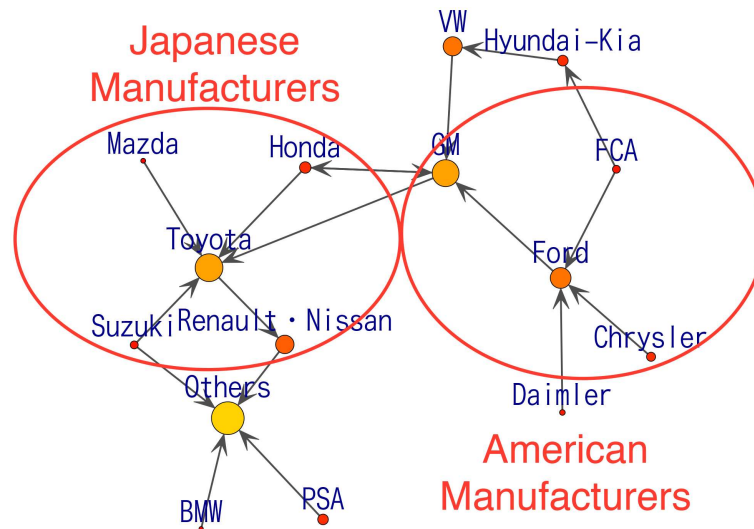


Fig 5. The estimated directed graph for “2007-1Q”.

doi:10.1371/journal.pone.0169981.g005

term (“2007-1Q”) of all records. The size of the circle at each node represents the market share of the corresponding manufacturer. The transition matrix is asymmetric. The (i, j) element of matrix G_t denoted by $(G_t)_{ij}$ is the transition probability from the i -th product to the j -th product in the time interval $(t - 1, t]$. We draw an arrow from the i -th node to the j -th node with a width proportional to the magnitude of $(G_t)_{ij}$ in the graph, representing the transition matrix at time t . In the directed graph, arrows connecting the nodes indicate that there are flows of sales share or flows of customers in the direction indicated by the arrows. Bi-directed arrows indicate that both connected nodes have in/out flows. The results show two distinct groups. One group includes American manufacturers such as Chrysler Group, Ford Group, GM Group, and the other group includes Japanese manufacturers such as Suzuki, Honda, TOYOTA Group, Mazda, and Renault-Nissan. Interestingly, there is a bi-directed arrow between GM and Honda, which are alliance companies. This phenomenon is presumably because the car dealer of each manufacturer recommends the cars of the alliance partner to customers. It is also possible that the same car dealer sells both GM and Honda cars.

Explanation of Social Events Inferred from the Transition Matrix

We present a more detailed analysis. First, within the fiscal year ending March 2008, TOYOTA Group has become the world’s top seller by beating GM Group in unit sales. We show the estimation results from “2007-1Q” to “2007-4Q” in Fig 6. From Fig 6(a), a direct flow from GM Group to TOYOTA Group is observed at 2007-1Q (red arrow). Then, for the next two quarterly periods shown in Fig 6(b) and 6(c), we infer that TOYOTA Group’s activity caused a reaction, and the graph shows an arrow from TOYOTA Group to GM Group (blue arrow). After this fluctuating behavior, at the fiscal year end in December shown in Fig 6(d), we do not see salient movement in consumers between those two manufacturers.

We next focus on 2009-1Q in Fig 7. In 2009, TOYOTA Group launched a massive recall and decreased its share to a large extent. This can be seen by comparing the size of the circles for TOYOTA Group in 2008-4Q shown in Fig 7(a) and in 2009-1Q shown in Fig 7(b). Comparing the graphs in 2008-4Q shown in Fig 7(a) and 2009-1Q shown in Fig 7(b), we see that from 2009, there is an arrow from TOYOTA Group to Others (red arrow), which indicates the defection of customers.

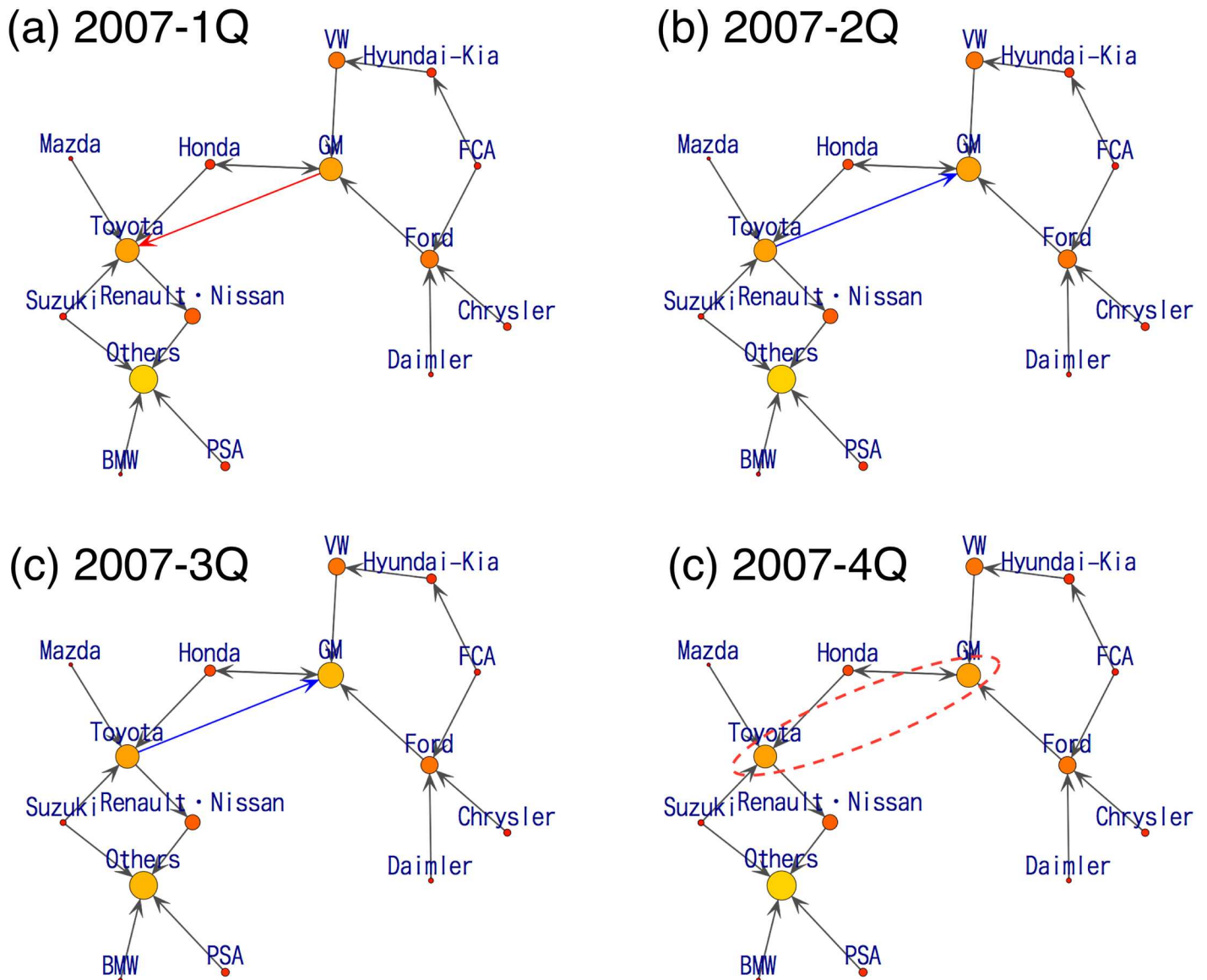


Fig 6. Visualization of consumers' flow estimated from a series of data from year 2007. There is a remarkable change in the direction and presence of the arrow between TOYOTA Group and GM Group.

doi:10.1371/journal.pone.0169981.g006

Finally, we focus on the year 2013 (Fig 8). In the second half of this year, VW Group beats GM Group in total sales amount to claim second position in the automobile industry. From Fig 8(a) and 8(b), the consumers' flow from VW Group to GM Group disappears in the 4th quarter in 2013, which indicates improvements in the brand image of VW compared to GM.

Conclusion

Summary of Contribution

In this paper, we considered the situation that manufacturers compete for limited consumers. By modeling the transition of consumers between different manufacturers using a Markov chain, we proposed a method to infer a sequence of transition matrices of consumers to

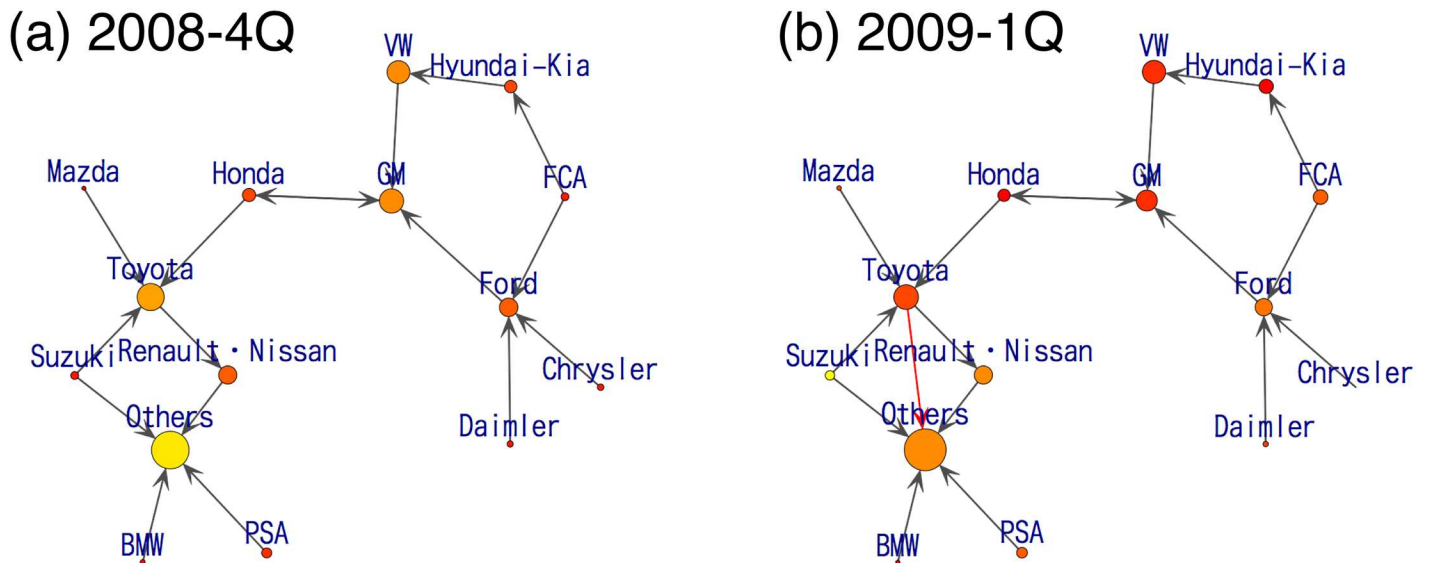


Fig 7. Visualization of estimated consumers' flow from a series of data in 2008-4Q and 2009-1Q.

doi:10.1371/journal.pone.0169981.g007

different manufacturers using a sequence of sales share data only. In the proposed model, the observed sales share data are identified with stationary probabilities of underlying Markov processes characterized by transition matrices. Assuming that the change in the structure, namely, a change in the transition matrices for consequent time is minor, we formulated the estimation problem of transition matrices as simple linear programming. The proposed method is applied to sales data for automobiles, and we obtained reasonable and socially explainable results. We believe that the results are significant in the sense that we can infer the flow of consumers only from sales share data. The results can be utilized for a market analysis or to develop a brand strategy with limited observations. For illustrative purposes, we considered the transition of consumers among manufacturers. However, the proposed method is

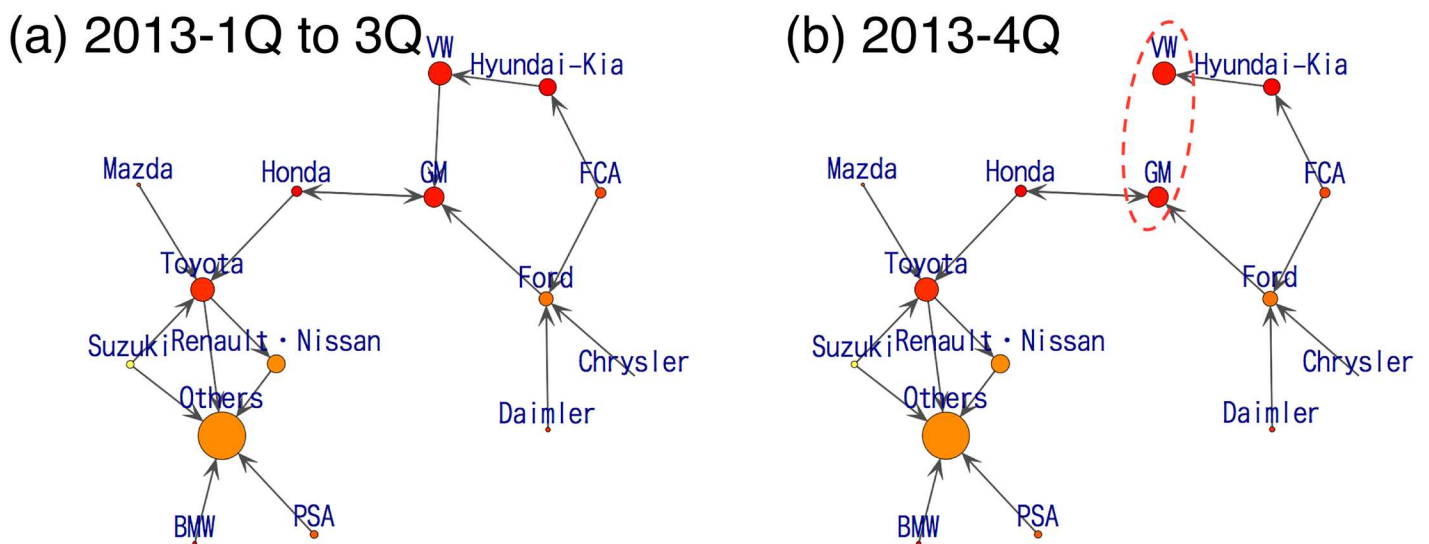


Fig 8. Visualization of estimated consumers' flow from a series of data in year 2013.

doi:10.1371/journal.pone.0169981.g008

applicable to more general situations with a fixed amount of *resources*, which is an abstraction of consumers, and nodes compete for finite resources through the edges.

Future Work

To estimate the transition matrices, we imposed necessary constraints so that the estimates should be transition matrices. Under these constraints, we minimized absolute difference between consequent matrices. There are other possibilities for optimizing objectives and constraints to improve the accuracy of estimation and interpretability. It would also be interesting to directly model the change in transition matrices with appropriate probability models.

Supporting Information

S1 Code. Program code and dataset to reproduce the results. Python code for the proposed method, and original dataset are available as a supporting information file. (ZIP)

Acknowledgments

The authors would like to express our special thanks to the editor and reviewers whose comments led to valuable improvements of the manuscripts.

Author Contributions

Conceptualization: NM SA HH.

Data curation: TC.

Formal analysis: NM SA HH.

Funding acquisition: NM SA HH.

Investigation: TC HH.

Methodology: NM SA HH.

Project administration: NM.

Resources: TC HH.

Software: TC HH.

Supervision: NM.

Validation: NM SA HH.

Visualization: TC HH.

Writing – original draft: TC HH SA NM.

References

1. Bondy JA, Murty USR. Graph theory. Graduate texts in mathematics. New York, London: Springer; 2007. OHX. Available from: <http://opac.inria.fr/record=b1123512>.
2. Gutierrez-Alcaraz G, Sheble GB. Electricity market price dynamics: Markov process analysis. 2004 International Conference on Probabilistic Methods Applied to Power Systems. 2004;p. 14–19.
3. Genon-Catalot V, Jeantheau T, Laredo C. Stochastic volatility models as hidden Markov models and statistical applications. *Bernoulli*. 2000; 6(6). doi: [10.2307/3318471](https://doi.org/10.2307/3318471)

4. Brown JH. *Macroecology*. The University of Chicago Press; 1995.
5. Bailey NTJ. *The elements of stochastic processes with applications to the natural sciences*. vol. 19. Wiley-Interscience; 1990.
6. Kleinberg JM. Authoritative sources in a hyperlinked environment. *Journal of the ACM (JACM)*. 1999; 46(5). doi: [10.1145/324133.324140](https://doi.org/10.1145/324133.324140)
7. Page L, Brin S, Motwani R, Winograd T. The PageRank citation ranking: Bringing order to the web. Stanford InfoLab; 1999. 1999–66.
8. Ng A, Zheng AX, Jordan MI. Link analysis, eigenvectors and stability. *Proceedings of the 17th international joint conference on Artificial intelligence*. 2001;.
9. Box GEP, Jenkins GM, Reinsel GC. *Time Series Analysis: Forecasting and Control*. vol. 3. John Wiley & Sons; 2013.
10. Fildes R. *Forecasting, Structural Time Series Models and the Kalman Filter*. vol. 8. Cambridge University Press; 1992. doi: [10.1016/0169-2070\(92\)90072-H](https://doi.org/10.1016/0169-2070(92)90072-H)
11. Wu Y, Hernández-Lobato JM, Ghahramani Z. Dynamic Covariance Models for Multivariate Financial Time Series. *Proceedings of the 30th International Conference on Machine Learning (ICML-13)*. 2013;.
12. Kummerfeld E, Danks D. Tracking time-varying graphical structure. *Advances in Neural Information Processing Systems 26 (NIPS 2013)*. 2013;p. 1–9.
13. Wright S. Correlation and causation. *Journal of agricultural research*. 1921; 20(7):557–585.
14. Pearl J. *Causality: Models, Reasoning, and Inference*. Cambridge University Press; 2001.
15. Granger CWJ. Investigating causal relations by econometric models and cross-spectral methods. *Econometrica: Journal of the Econometric Society*. 1969;p. 424–438. doi: [10.2307/1912791](https://doi.org/10.2307/1912791)
16. Friedman J, Hastie T, Tibshirani R. Sparse inverse covariance estimation with the graphical lasso. *Biostatistics*. 2008 Jul; 9(3):432–441. doi: [10.1093/biostatistics/kxm045](https://doi.org/10.1093/biostatistics/kxm045) PMID: [18079126](https://pubmed.ncbi.nlm.nih.gov/18079126/)
17. Xuan X, Murphy K. Modeling changing dependency structure in multivariate time series. *Proceedings of the 24th international conference on Machine learning*. 2007;p. 1055–1062.
18. Ide T, Lozano AC, Abe N, Liu Y. Proximity-based anomaly detection using sparse structure learning. *SDM*. 2009;.
19. Ahmed A, Xing EP. Recovering time-varying networks of dependencies in social and biological studies. *Proceedings of the National Academy of Sciences of the United States of America*. 2009 jul; 106(29):11878–11883. doi: [10.1073/pnas.0901910106](https://doi.org/10.1073/pnas.0901910106) PMID: [19570995](https://pubmed.ncbi.nlm.nih.gov/19570995/)
20. Langville N, Meyer D. Deeper inside PageRank. *Internet Mathematics*. 2004; p. 335–380. doi: [10.1080/15427951.2004.10129091](https://doi.org/10.1080/15427951.2004.10129091)
21. S George P H, Smith H. Markov Chains Applied to Marketing. *Journal of Marketing Research*. 1964; p. 50–55.
22. Haveliwala T H, Kamvar S D. The second eigenvalue of the google matrix Stanford University Technical Report, No. 1 2003.
23. Tibshirani R. Regression Shrinkage and Selection Via the Lasso. *Journal of the Royal Statistical Society, Series B*. 1996; 58:267–288.
24. Elad M. *Sparse and Redundant Representations: From Theory to Applications in Signal and Image Processing*. Springer; 2010. doi: [10.1007/978-1-4419-7011-4](https://doi.org/10.1007/978-1-4419-7011-4)
25. Tibshirani R, Saunders M, Rosset S, Zhu J, Knight K. Sparsity and smoothness via the fused lasso. *J R Stat Soc Ser B Stat Methodol*. 2005; 67(1):91–108. doi: [10.1111/j.1467-9868.2005.00490.x](https://doi.org/10.1111/j.1467-9868.2005.00490.x)
26. Dantzig GB, Thapa MN. *Linear Programming. 2., Theory and extensions*. Springer series in operations research. New York: Springer; 2003.
27. Makhorin A. GLPK (GNU Linear Programming Kit);. Available at <http://www.gnu.org/software/glpk/glpk.html>.