2020 5th IEEE International Conference on Big Data Analytics

Testing the Dependence between Network Positions and National Attributes in Countries Interaction Network

Yu Song
School of Software Engineering
Huazhong University of Science and Technology
Wuhan, China
e-mail: javasong@foxmail.com

Dehong Qiu
School of Software Engineering
Huazhong University of Science and Technology
Wuhan, China
e-mail: qiudehong@163.com

Fuxiang Yan
School of Software Engineering
Huazhong University of Science and Technology
Wuhan, China
e-mail: jackacmlove@gmail.com

Abstract—**Exploring the relationship between the positions of countries in the countries interaction network and their development status is of great significance in world system studies. To characterize the complex interactions between countries, network methodologies have been extensively employed. The methods typically used in the field include blockmodel analysis and statistical models, which have the deficiencies of instability, dependency on specific assumptions, and requirement of domain knowledge. Enlightened by the network embedding methods, we propose a novel framework combining network embedding and correlation analysis for addressing the problem in a model-free data-driven way, and provide a quantitative test statistic representing the magnitude of the correlation. We test our method on a real world dataset containing countries' trade and political relations as well as several indices of development levels. Experiment results derived from our method are in consistency with previous findings via other approaches, and provide extra information about the pattern of the relationship.**

Keywords-international network; network embedding; dependence testing; network analysis; world system

I. INTRODUCTION

The world has long been a complicated system with members interacting with each other. The interactive relations among countries include commodity trade, political support, diplomatic ties and force intervention. The structure of the relationship is represented by networks, with each node representing a country, and the presence of an edge indicating the existence of the relation. Identifying the relationship between the position of a country in the interaction network and its attributes is of crucial significance. For example, dose the dependency structure of a country affect its economic growth rate [1][2]? Is there an association between countries' positions and their population health levels [3]? The world system theory provides a hypothesis for the question.

The world system theory was first proposed in the 1970s [4][5], which divides the world into core countries, semi-periphery countries and periphery countries. The theory suggests that the dependency structure within the global system, i.e., the roles that countries play in it, is crucial for understanding various dimensions of development within countries. Since then, many researchers have focused on evaluating the validity of these theories [6][7]. The solution to the problem is not only important to the study of the world system, but also provides useful advice to decision makers and affects every aspect of life.

A lot of efforts have been made on solving the problem, among which the most canonical approaches are the blockmodel analysis and statistical models [8][9]. Such methods rely on hierarchical clustering or parameter estimating to obtain representations for countries' network positions, followed by statistical methods such as regression analysis and independence testing based on the obtained position representations. However, these methods suffer from several issues. On the one hand, the positions derived by blockmodels are qualitative indices, which limit the power of post statistical analysis. On the other hand, statistical models are highly dependent on specific assumptions, which is difficult to guarantee.

To overcome the drawbacks of these methods, we propose a novel framework combining network embedding and statistical correlation analysis. We first embed the nodes (representing countries) in the network to a Euclidean space, then apply traditional statistical methods to study the correlation between the obtained node embeddings and nodal attributes. Our method is model-free and data-driven, and provides a direct quantified result representing the magnitude of the correlation.

The rest of the paper is organized as follows. Section 2 reviews important literatures concerning the problem. Section 3 presents the basic framework of our method, while section 4 covers the technical details. Experiment setting and results can be found in section 5, where we test our method on multiple real world countries interaction networks. We conclude the paper in section 6, discussing the limitations of our method, and point out future directions.

II. RELATED WORKS

As discussed above, the problem of correlation analysis between the positions of countries and their attributes is twofold: (1) identify the position of each country in the international interaction network and (2) test the dependence between positions and national attributes. In this section, We review two canonical approaches to tackle the problem, i.e., blockmodel analysis and statistical models. In addition, we introduce the network embedding method which has achieved brilliant results in multiple fields, and discuss how it can be of help in this issue.

A. *Blockmodel Analysis*

Block model analysis, among other approaches, is one of the most preferred methods for analyzing network positions [10]. The process of blockmodel analysis can be described as repeatedly permuting the rows and columns of the adjacency matrix, until the new matrix exhibits a distinct structure of 0-blocks and 1-blocks, with 1-blocks containing a high density of edges and 0-blocks on the contrary. Nodes in the same block therefore share the same relation with other blocks. Following this procedure, statistical analysis is employed on nodal attributes based on the blocks discovered by the method. For example, Snyder and Kick [1] used a hierarchical clustering algorithm CONCOR to partition countries based on four types of international relations, the result of which provides strong evidence for a core–semiperiphery–periphery structure. They then applied regression analysis to test the impact of structural positions identified by the block model on nations' economic growth rate. The outcome shows a strong dependence between the two variables and is in great consistency with the world system theories. Moore et al. [3] applied blockmodeling techniques to the international trade data circa 2000, and derived a similar core-semiperiphery-periphery structure of countries. Regression analysis show that globalization and trade may have disproportionate effects on national institutions, policies, and population health, depending on the positions of countries. York et al. [11] examined the effects of nations' positions in the world system on women's status within the nations. The positions were coded into dummy variables, and the results of OLS regression indicate that women's status is associated with national links to the global economy, and gender equality leads nations to be less dependent on the world economic system.

Despite the decent performance of blockmodel analysis, its deficiencies are obvious. First, due to the qualitative nature of the position identification procedure, this mode of analysis is indeed a qualitative way to study the correlation, rather than quantitative. Second, the partition results from the block model are highly subject to the specific algorithm that is being used, making it hard to choose which one to use.

B. *Statistical Models*

To study the correlation between network positions and nodal attributes, model-based approaches have been extensively used. These approaches involve model construction and parameter estimating, and conclusions are drawn from the estimated parameters after the fitting procedure. For instance, Pattison and Elliott generalized the $p*$ class of models [12] for social network data to predict individual attributes from network relations. They fitted the model with a social network dataset which contains the social relationship between 14 trainees together with the attitudes (disagreement, neutrality, and agreement) of each trainee toward a list of statements. The modeling approach they developed provided some evidence that people indeed tend to be affected by their partners in the social network. Fosdick and Hoff [13] developed a latent variable model to obtain node-specific low-dimensional representations of network structure, then introduced a testing procedure for linear dependence between nodal network factors and attributes based on the assumption of multivariate normal distribution. They tested the model on the National Longitudinal Study of Adolescent Health data consisting of 389 high school students, and the result of hypothesis testing showed that students' friendships are related to their health behaviors and grade point average (GPA) to a considerable degree. Model-based methods as such have succeeded in quantifying the magnitude of the correlation, however, they also have some limitations. For example, like any model based methods, the effectiveness of statistical model analysis depends on the consistency of real data and model assumptions, which is difficult to guarantee. Besides, the construction of the model and the interpretation of model results require a vast amount of knowledge in the relevant fields.

C. *Network Embedding*

Studying relational data is one of the core problems in many scientific fields including biology, chemistry and social science. Relational data often takes the form of networks, which are usually represented by adjacency matrices. While this representation has the advantages of simplicity in form and ease of use, it suffers from issues such as high computational complexity, sparsity of data, and most of all, inapplicability of most statistical methods that take independent feature vectors as input. To tackle these challenges, substantial effort has been made to derive low dimensional vector representations, i.e., embeddings, for network nodes that capture the complex dependency structure within the network. Recent years have seen great progress in the network embedding field [14][15], and the derived embeddings have been applied to tasks including node classification [16][17], node clustering [18], network visualization [19] and link prediction [20][21].

However, the embedding method has not been widely used in tasks of correlation analysis, which deserve more attention. In 2018, Lee et al. introduced a novel method called the diffusion correlation to test the dependence between network positions and nodal attributes [22]. They used a network embedding algorithm called the

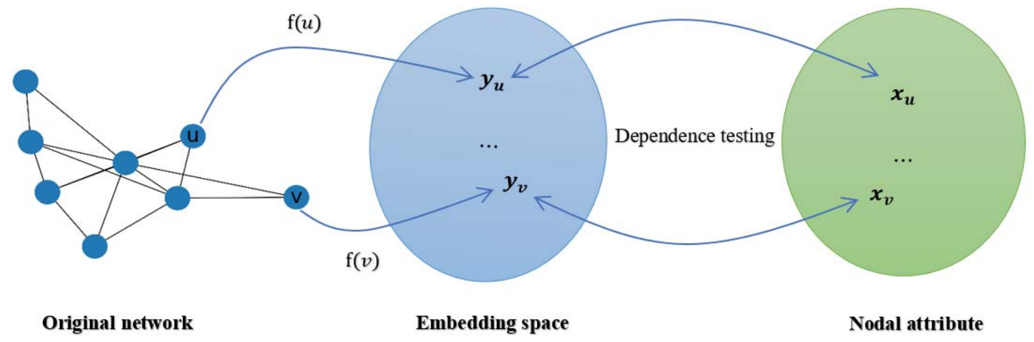


Figure 1. Framework of correlation analysis between network position and nodal attribute

diffusion map to obtain structural representations of nodes in a low-dimensional Euclidean space, followed by traditional statistical methods for independence testing. They test their method on a biological network, demonstrating the existence of the relationship between the physical positions and network positions of neurons in the organism. To the best of our knowledge, this is the first attempt to combine network embedding and statistical correlation analysis. Nonetheless, embeddings generated by the diffusion map represent only the abstract connectivity between nodes and does not contain concrete meanings. Our method differs from the diffusion correlation in the following aspects: (1) we adopt a different embedding algorithm from the diffusion map, (2) we employ two similarity measures to obtain node embeddings with different meanings and (3) with the concrete meanings of node embeddings, the interpretability of experiment results is greatly enhanced.

III. PROBLEM DEFINITION&TECHNIQUE FRAMEWORK

The complex interactions among countries are represented by networks. A network $G(V,E)$ is composed of a set of nodes $V = \{v_1, \cdots, v_n\}$ representing countries and an edge set $E = \{e_{ij}\}_{i,j=1}^{n}$ representing the relationship between them. The adjacency matrix A associated with the network is a matrix of $R^{n\times n}$ consisting of 0s and 1s, with 1 indicating the presence of the relation and 0 otherwise . The network can be directed or undirected, depending on the property of the relationship. We denote the nodal attributes of the network as a matrix $X \in R^{n\times p}$ and each row of X represents the attributes of the corresponding node.

To characterize the dependency structure within the network, we introduce the network embedding method. Given a network $G = (V,E)$, a network embedding is a mapping $f: v_i \to \boldsymbol{y_i} \in R^d$, $\forall i \in [n]$ such that $d \ll |V|$, and $Y = [\boldsymbol{y_i}]$ is the embedding matrix of network G. A proper mapping f will preserve the relations between nodes from the original network.

After obtaining the node embeddings, we test the dependence between $\boldsymbol{x}$ and $\boldsymbol{y}$ by computing the correlation coefficient between them. The framework is shown in Fig. 1, and the details are presented in the following section.

IV. DETAILS OF OUR METHOD

The network embedding algorithm we adopt is based on the Graph Factorization (GF) [23], with some modification to better suit the task of dependence testing. We employ the distance correlation to measure the dependence between network positions and nodal attributes, mainly for its capability of capturing high-dimensional non-linear relationship. The description of the framework is given in Table I.

A. *Graph Factorization*

Compared with other network embedding methods, Graph Factorization has the advantages of few hyper-parameters, good scalability, and most of all, the convenience of introducing multiple similarity measures. We provide the derivation process of the algorithm here, and discuss its property.

In order to generate embeddings that preserve the properties from the original network, a common practice adopted by embedding algorithms is to optimize a loss function. The loss function aims to maintain the consistency of the similarity between nodes in the original network and the similarity between nodes in the embedding space. A classic loss function takes the form of

$$L(Y) = \sum_{i,j=1}^{n} \left(A_{ij} - \langle Y_i, Y_j \rangle\right)^2 . \quad (1)$$

Equation (1) uses dot-product as the similarity measure between node embeddings. It attempts to minimize the difference between the embedding similarity and the corresponding entry in the adjacency matrix. From another perspective, it aims to reconstruct the adjacency matrix A by $A = Y^T Y$. However, as can be seen in the formula, the time complexity of the algorithm is $O(|V|^2)$, making it not scalable to large datasets. To relieve the computational burden, Ahmed et al. made a small modification to the formula [23], and the new loss function looks like

$$L(Y,\lambda) = \frac{1}{2} \sum_{(i,j)\in E} (A_{ij} - \langle Y_i, Y_j \rangle)^2 + \frac{\lambda}{2} \sum_i \|Y_i\|^2 . \quad (2)$$

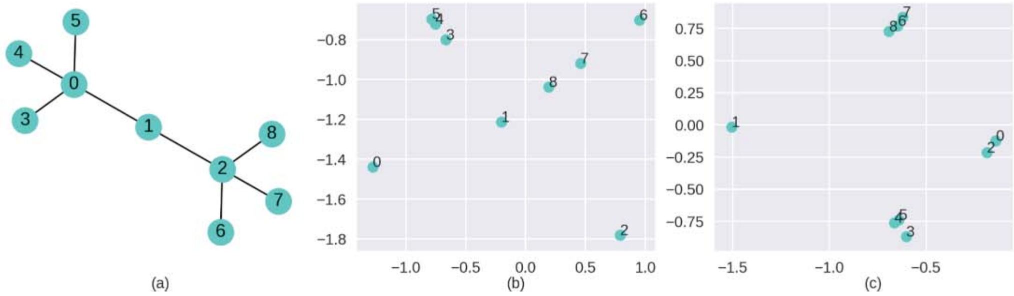


Figure 2. Example motif and node embeddings visualization

Equation (2) is called the Graph Factorization. It reduces the time complexity from $O(|V|^2)$ to $O(|E|)$ by only summing over non-zero elements in the adjacency matrix. Thus, it is an approximation of the original adjacency matrix. The right part of the formula serves as a regularization term to prevent overfitting, and λ is the regularization coefficient.

In the above methods, the adjacency matrix A is used to guide the optimization process. The adjacency matrix can be viewed as a measure of similarity between nodes, which models the direct connections within the network. However, there exists a lot of dependency structures that cannot be described by direct connections. Here we introduce a widely used similarity measure between nodes which is called structural equivalence [24]. Several algorithms for calculating structural equivalence have been proposed. Despite the multifarious formulas, they all aim to describe the similarity of the neighborhoods of a pair of nodes. In this paper, we implement a basic version of structural equivalence, i.e., the Jaccard similarity, as shown in (3), where $N(\cdot)$ means the neighborhood of a node.

$$S(v_i, v_j) = \frac{|N(v_i) \cap N(v_j)|}{|N(v_i) \cup N(v_j)|} \tag{3}$$

Combining the loss function in (2) and the Jaccard structural equivalence, we now have another version of the Graph Factorization:

$$L(Y, \lambda) = \frac{1}{2} \sum_{(i,j)\in E} (S_{ij} - \langle Y_i, Y_j \rangle)^2 + \frac{\lambda}{2} \sum_i \|Y_i\|^2, \tag{4}$$

where S means the similarity matrix. The difference between direct connections and structural equivalence is shown in Fig. 2. Fig. 2 (a) shows an example motif, while Fig. 2 (b) and Fig. 2 (c) plot the visualization of the 2-dimensional node embeddings derived via the adjacency matrix and the Jaccard similarity matrix respectively. Node 0 and node 2 are far away from each other in Fig. 2 (b) because they are not directly connected, whereas in Fig. 2 (c), the two nodes are pretty close, for they share a common neighbor, i.e., node 1. The example shows that node embeddings generated via different similarity measures carry different meanings; thus, they can be used to identify different correlation patterns between node positions and attributes.

The optimization of GF can be solved by gradient decent. The gradient of L in (4) with respect to the row i of Y is given by

$$\frac{\partial L}{\partial Y_i} = -\sum_{j\in N(i)} (S_{ij} - \langle Y_i, Y_j \rangle) Y_j + \lambda Y_i. \tag{5}$$

For a pair of node $(i, j) \in E$, (5) equals to

$$-(S_{ij} - \langle Y_i, Y_j \rangle) Y_j - \lambda Y_i. \tag{6}$$

Thus, the update step for each node embedding Y_i is

$$Y_i^{t+1} \leftarrow Y_i^t + \eta[(S_{ij} - \langle Y_i, Y_j \rangle) Y_j + \lambda Y_i]. \tag{7}$$

Till now, the formulation and optimization of the loss function have been solved. In this article, we implement both (2) and (4) to derive embeddings that preserve different similarities.

B. Distance Correlation

Testing the dependence between two random variables is one of the core problems in statistics. The most famous approach concerning this problem is the Pearson's correlation coefficient [25], which reports the extent of linear relationship in the given sample. Another widely used method, i.e., canonical correlation analysis (CCA) [26], can be viewed as the enhanced version of Pearson's correlation coefficient, with the capability of determining linear correlation in high-dimensional data. However, capturing only linear relationship is not enough in many scenarios, including network analysis. To tackle the problem, Szekely et al. proposed a distance based correlation measure for identification of high dimensional non-linear association between two random vectors called the distance correlation [27].

The distance covariance is defined in equation (8). $\boldsymbol{x}$ and $\boldsymbol{y}$ are random vectors, and $D(\cdot,\cdot)$ means Euclidean distance.

$$dCov(X, Y) = \frac{1}{n^2} \sum_{i=1}^{n} \sum_{j=1}^{n} D(\boldsymbol{x_i}, \boldsymbol{x_j}) \cdot D(\boldsymbol{y_i}, \boldsymbol{y_j}). \tag{8}$$

The distance correlation is obtained by normalizing the distance covariance to $[0,1]$:

$$dCor(X, Y) = \frac{dCov(X, Y)}{\sqrt{dCov(X, X) \cdot dCov(Y, Y)}}. \tag{9}$$

TABLE I. DESCRIPTION OF OUR FRAMEWORK

Framework: **Correlation analysis between nodal positions and attributes**
Input Adjacency matrix $A \in R^{n\times n}$ Attribute matrix $X \in R^{n\times p}$ Embedding dimension d
Process 1. Compute similarity matrix $S \in R^{n\times n}$ from A via specified similarity measure 2. Initialize embedding matrix $Y \in R^{n\times d}$ 3. Optimize loss function to obtain Y 4. Compute distance correlation coefficient t between X and Y 5. Apply standard permutation test to derive significance level p
Output Distance correlation coefficient t Significance level p

TABLE II. MEAN ATTRIBUTE WITHIN BLOCKS ON COUNTRIES TRADE NETWORK

National Attribute	Block Number					
	B1	*B2*	*B3*	*B4*	*B5*	*B6*
Education level	90.00	57.00	51.14	43.67	47.33	14.33
Energy consumption	7217.67	2615.40	2892.29	791.00	1265.67	200.00
Population growth rate	0.73	1.30	1.60	3.17	3.13	2.70

The distance correlation coefficient measures the extent of the co-varying of the pairwise distances of samples from X and Y. It takes the value of zero if and only if there is no dependence of any kind between the two random vectors. In other words, a zero correlation coefficient implies independence. With the distance correlation, we are able to test various kinds of dependencies within networks.

V. EXPERIMENT RESULTS

We employed our framework on a real world network dataset of international trade flows and diplomatic ties. We analyzed the correlation between countries' network positions and their development status, and obtained some information about the dependency patterns within the network, which provides insight into the property of the relationship.

A. Data Set Description

The dataset we used to test our method is called countries trade networks, collected by Wasserman et al. in 1994 for educational purpose [28]. The dataset contains 24 countries, chosen to represent a wide range of interesting economical features and to span the categories of existing world system and development levels. Five relations were measured on the countries, as shown in the following:

Relations on the 24 countries
1. Imports of raw materials
2. Imports of basic manufactured goods
3. Imports of food and live animals
4. Imports of mineral fuels
5. Diplomatic exchange

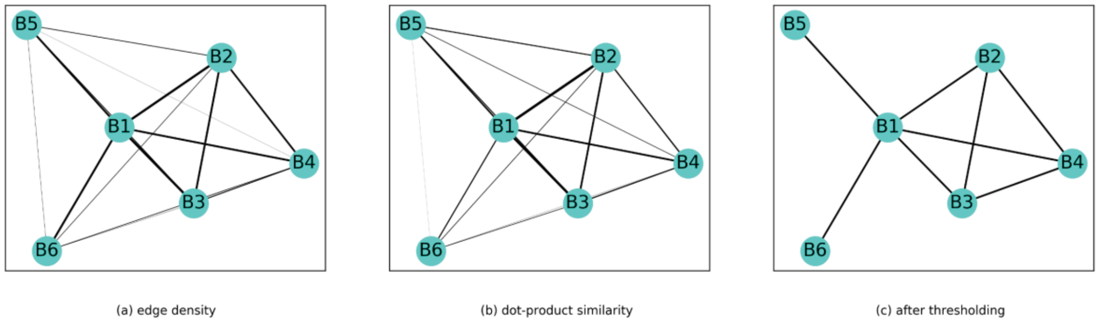
The first four relations are economic, and the last one is political. Each relation is represented by a 24×24 binary adjacency matrix. Besides relations among countries, four national attributes were collected, aiming to reflect the economic and development status of the countries. The attributes are:

Attributes of the 24 countries
1. Annual population growth rate between 1970 and 1981
2. Annual growth of GNP per capita from 1970 to 1981
3. Secondary school enrollment ratio in 1980
4. Energy consumption per capita in 1980 (measured in kilo coal equivalent)

Wasserman et al. employed a hierarchical clustering algorithm in SYSTAT [29] on three relations among the 24 countries: imports of basic manufactured goods, imports of raw materials and diplomatic ties. The algorithm produced a six-block model representing six types of countries. The blocks and the corresponding members are listed in the following.

Description of the six-block model
B1: Japan, United Kingdom, United States
B2: China, Czechoslovakia, Indonesia, Spain, Yugoslavia
B3: Argentina, Brazil, Finland, New Zealand, Pakistan, Switzerland, Thailand
B4: Algeria, Egypt, Syria
B5: Ecuador, Honduras, Israel
B6: Ethiopia, Liberia, Madagascar

They calculated the block means of the national attributes. They argued that these variables should differ systematically across the positions in the model, if the nation's dependency structure dose relate to their development status. The results of the experiment are shown in Table Ⅱ. From the table, we can see that the partition of countries is arranged in a sort that basically characterizes the development levels of countries. Countries in B1 block are with no doubt tier one countries in the world by that time, e.g., the United States and the United Kingdom. The common characteristics of these countries are low population growth rate, high education levels, and extremely high energy consumptions, indicating the mature industrial systems and high living standards. On the other side, members within B6 block are all undeveloped African countries, with low industrial levels as well as high population growth rate. With the results from the six-block model, they concluded that there does exist a



(a) edge density (b) dot-product similarity (c) after thresholding

Figure 3. Associations between blocks with respect to edge density and node embedding similarity

dependence between countrys' positions and development status.

B. *Our Results*

In order to employ our framework of correlation analysis, we calculated the similarity matrices between nodes based on both direct connections and structural equivalence. When calculating structural equivalence, incoming and outgoing neighbors are considered separately. After this, we fed the similarity matrices to the embedding algorithm GF and obtained a 2-dimensional vector representation for each node in the network.

The validity of the embedding method can be illustrated by a simple example. Fig. 3 (a) shows the edge densities between the six blocks of countries derived by hierarchical clustering as mentioned above. The thickness of lines indicates the magnitudes of the associations between blocks. In Fig. 3 (b), the thickness of lines represents the mean dot-product similarity between node embeddings across the same six blocks. By setting proper thresholds, the two methods exhibit perfect consistency (c).

We computed the distance correlation coefficient between the obtained node embeddings and nodal attributes, followed by a standard permutation test. In order to maintain consistency with the previous method, we did not perform additional processing on the attribute data. To remove the randomness from GF, the procedure was repeated for 50 times, and the mean test statistic and the number of times the independence assumption is accepted are reported in Table Ⅲ, Table IV and Table V. As we can see, most of the assumptions have not been accepted once, indicating the relations are statistically significant.

For example, in the raw materials trade network (Table Ⅲ), all associations except the one between direct connections and energy consumptions are shown to be significant, suggesting the interplay between raw materials trade and national development. The exception may be due to the huge standard deviation of the energy consumption values, which introduces noises to the computing of distance correlation.

In the diplomat exchange network (Table IV), the direct connections between countries are proven to be not significantly related to either of the three attributes (first row), whereas the similarity between countries' outgoing neighbors and their attributes are shown to be significantly relevant (last row). This phenomenon reveals that the diplomatic ties between countries are hardly related to the their economic and industrial levels, however, if two countries share a lot of common allies, they are very likely to resemble each other.

In the third column of the manufactured goods trade network (Table V), we observe that although the correlations between population growth rate and the three types of dependency structures are all significant, the magnitude of them varies. Countries' common export destinations, among others, exhibit the strongest correlation with population growth rate. Fig. 4 shows the scatter plots of the pairwise Euclidean distances of nations' population growth rate and the three types of structural embeddings, where we can see a clear increment of correlation from (a) to (c). A possible explanation for the strong correlation in Fig. 4 (c) may be that if two countries have similar export destinations, they are highly likely to have similar positions in the world industry chain, which determine the countries' development levels to a considerable degree. Remember, population growth rate is an important indicator of development status.

TABLE III. Mean Correlation Coefficient on Countries Trade Network of Raw Materials

Similarity Measure	Nodal Attribute		
	Education level	*Energy consumption*	*Population growth*
Direct connection	0.544, 0	**0.488, 30**	0.550, 0
Incoming neighbor	0.649, 0	0.655, 0	0.699, 0
Outgoing neighbor	0.684, 0	0.648, 0	0.750, 0

TABLE IV. Mean Correlation Coefficient on Countries Diplomat Exchange Network

Similarity Measure	Nodal Attribute		
	Education level	*Energy consumption*	*Population growth*
Direct connection	**0.411, 45**	**0.373, 47**	**0.379, 49**
Incoming neighbor	0.492, 2	0.410, 50	0.590, 0
Outgoing neighbor	**0.544, 0**	**0.491, 0**	**0.647, 0**

TABLE V. Mean Correlation Coefficient on Countries Trade Network of Basic Manufactured Goods

Similarity Measure	Nodal Attribute		
	Education level	*Energy consumption*	*Population growth*
Direct connection	0.596, 0	0.552, 0	**0.659, 0**
Incoming neighbor	0.692, 0	0.684, 0	**0.785, 0**
Outgoing neighbor	0.610, 0	0.681, 0	**0.841, 0**

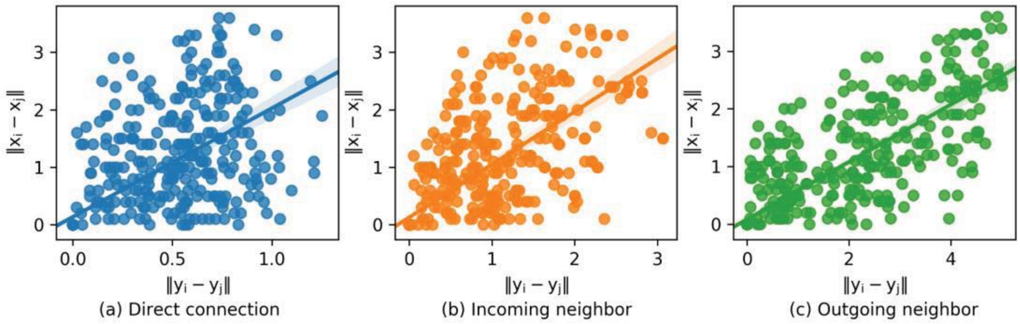


Figure 4. Paired Euclidean distances for countries' positions and population growth rates

VI. CONCLUSION

We proposed a novel framework for analyzing the correlation between network positions and nodal attributes that produces quantitative measurements of positions via network embedding and uses distance correlation coefficient to capture the complex non-linear relationship between nodes. Experiment results on the countries interaction network show that according to the definition of position, countries' positions in the network have varying degrees of correlation with their level of development, which is in great consistency with the world system theories. Besides, our method gives a specific value for the magnitude of the correlation, and provide insight into the dependency patterns.

Unlike the diffusion correlation, our method implemented a different embedding algorithm called Graph Factorization, which generates embeddings with concrete meanings.

While the method presented here has the advantages of being quantitative and good interpretability, it also has some limitations. On the one hand, the similarity measures we used are too simple to characterize the numerous kinds of dependency structures within networks; on the other hand, the embedding method GF may have trouble optimizing the loss function when the similarity matrix is complicated. In addition, a key problem here is the lack of theoretical proof for the consistency of network embedding and the distance correlation coefficient.

We view our method as a compliment to the traditional approaches, rather than a substitute for them. By jointly applying our method and traditional approaches such as blockmodeling, we can obtain a more comprehensive understanding about the network. For the next step, we plan to (1) prove the theoretical correctness of the combination of network embedding and correlation coefficient, and (2) develop more advanced embedding algorithms to capture the various kinds of relations within networks.

ACKNOWLEDGEMENT

The authors would like to thank their colleagues and the anonymous reviewers for their helpful comments and suggestions. Part of the research was supported by the National Natural Science Foundation of China 61873100.

REFERENCES

[1] Snyder, David, and Edward L. Kick, "Structural position in the world system and economic growth, 1955-1970: A multiple-network analysis of transnational interactions," *American journal of Sociology*, vol. 84, no. 5, pp. 1096-1126, 1979.

[2] Smith, David A., and Douglas R. White, "Structure and dynamics of the global economy: network analysis of international trade 1965–1980," *Social forces*, vol. 70, no. 4, pp. 857-893, 1992.

[3] Moore, Spencer, Ana C. Teixeira, and Alan Shiell, "The health of nations in a global context: trade, global stratification, and infant mortality rates," *Social science & medicine*, vol. 63, no. 1, pp. 165-178, 2006.

[4] Wallerstein and Immanuel, "The rise and future demise of the world capitalist system: Concepts for comparative analysis," *Comparative studies in society and history*, vol. 16, no. 4, pp. 387-415, 1974.

[5] Wallerstein, Immanuel Maurice, and Senior Researcher Immanuel Wallerstein, *World-systems analysis: An introduction*, Duke University Press, 2004.

[6] DiPrete, Thomas A., and Gregory M. Eirich, "Cumulative advantage as a mechanism for inequality: A review of theoretical and empirical developments," *Annu. Rev. Sociol*, vol. 32, pp. 271-297, 2006.

[7] Rossoni, Luciano, Márcio Luiz Marietto, and Wesley Vieira da Silva, "The effect of positioning in the world economic system on economic and social development: A relational approach to services," *African Journal of Business Management*, vol. 5, no. 20, pp. 8008-8024, 2011.

[8] White, Harrison C., Scott A. Boorman, and Ronald L. Breiger, "Social structure from multiple networks. I. Blockmodels of roles and positions," *American journal of sociology*, vol. 81, no. 4, pp. 730-780, 1976.

[9] Robins, Garry, Peter Elliott, and Philippa Pattison, "Network models for social selection processes," *Social networks*, vol. 23, no. 1, pp. 1-30, 2001.

[10] Arabie, Phipps, Scott A. Boorman, and Paul R. Levitt, "Constructing blockmodels: How and why," *Journal of mathematical psychology*, vol. 17, no. 1, pp. 21-63, 1978.

[11] York, Richard, and Christina Ergas, "Women's Status and World-System Position: An Exploratory Analysis," *Journal of World-Systems Research*, vol. 17, no. 1, pp.147-164, 2011.

[12] Robins, Garry, Philippa Pattison, and Peter Elliott, "Network models for social influence processes," *Psychometrika*, vol. 66, no. 2, pp. 161-189, 2001.

[13] Fosdick, Bailey K., and Peter D. Hoff, "Testing and modeling dependencies between a network and nodal attributes," *Journal of the American Statistical Association*, vol. 110, no. 511, pp. 1047-1056, 2015.

[14] Goyal, Palash, and Emilio Ferrara, "Graph embedding techniques, applications, and performance: A survey," *Knowledge-Based Systems*, vol. 151, pp. 78-97, 2018.

[15] Cui, Peng, Xiao Wang, Jian Pei, and Wenwu Zhu, "A survey on network embedding," *IEEE Transactions on Knowledge and Data Engineering*, vol. 31, no. 5, pp. 833-852, 2018.

[16] Huang, Xiao, Jundong Li, and Xia Hu, "Label informed attributed network embedding," in *Proceedings of the Tenth ACM International Conference on Web Search and Data Mining*, pp. 731-739, ACM, 2017.

[17] Dai, Quanyu, Qiang Li, Jian Tang, and Dan Wang, "Adversarial network embedding," in *Thirty-Second AAAI Conference on Artificial Intelligence*, 2018.

[18] Wang, Xiao, Peng Cui, Jing Wang, Jian Pei, Wenwu Zhu, and Shiqiang Yang, "Community preserving network embedding," in *Thirty-First AAAI Conference on Artificial Intelligence*, 2017.

[19] Wang, Daixin, Peng Cui, and Wenwu Zhu, "Structural deep network embedding," in *Proceedings of the 22nd ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 1225-1234, ACM, 2016.

[20] Wei-Chin Hung, Hung-Wei Lin, Yu-Chung Tsao, and Wei-Guang Teng, "Predicting Cooperation Relationships in Heterogeneous Movie Networks," *International Journal of Machine Learning and Computing*, vol. 4, no. 5, pp. 405-410, 2014.

[21] Hui Li, Chunmei Liu, and Legand Burge, "Predicting Protein-Protein Interactions Based on PPI Networks," *International Journal of Machine Learning and Computing*, vol.2, no. 6, pp. 794-797, 2012.

[22] Lee, Youjin, Cencheng Shen, Carey E. Priebe, and Joshua T. Vogelstein, "Network dependence testing via diffusion maps and distance-based correlations," *Biometrika*, vol. 106, no. 4, pp. 857-873, 2019.

[23] Ahmed, Amr, Nino Shervashidze, Shravan Narayanamurthy, Vanja Josifovski, and Alexander J. Smola, "Distributed large-scale natural graph factorization," in *Proceedings of the 22nd international conference on World Wide Web*, pp. 37-48, ACM, 2013.

[24] Lorrain, Francois, and Harrison C. White, "Structural equivalence of individuals in social networks," *The Journal of mathematical sociology*, vol. 1, no. 1, pp. 49-80, 1971.

[25] Pearson, Karl, "VII. Note on regression and inheritance in the case of two parents," in *Proceedings of the royal society of London*, vol. 58, pp. 240-242, 1895.

[26] Hotelling and Harold, "Relations between two sets of variates," in *Breakthroughs in statistics*, pp. 162-190, Springer, New York, NY, 1992.

[27] Székely, Gábor J., Maria L. Rizzo, and Nail K. Bakirov, "Measuring and testing dependence by correlation of distances," *The annals of statistics*, vol. 35, no. 6, pp. 2769-2794, 2007.

[28] Wasserman, Stanley, and Katherine Faust, *Social network analysis: Methods and applications*, vol. 8, Cambridge University Press, 1994.

[29] Wilkinson and Leland, *SYSTAT: the system for statistics*, Evanston, 1990.