On a Typology of Nodes and Its Applications in Network Analysis

Vladimir Matveenko

National Research University Higher School of Economics, St. Petersburg, Russia,
vmatveenko@hse.ru,
https://www.hse.ru/en/org/persons/202791

Abstract. Commonly in network analysis a graph (network) is represented by its adjacency matrix, and the latter may have an enormous order. We show that in many situations (generalizing the case of regular graph) a much smaller matrix (referred as type adjacency matrix) may be used instead. We introduce concepts of the types of nodes and of the type adjacency matrix, study properties of the latter and demonstrate some of its applications in social and economic network analysis. In particular, we consider centrality measures in undirected networks and dynamic patterns in a development model based on the structure of optimal paths in directed weighted networks.

Keywords: graph, network analysis, adjacency matrix, types of nodes, centrality, optimal paths

1   Introduction

Commonly in network analysis and its economic and social applications a graph (network) is represented by its adjacency matrix. A drawback of the adjacency matrix is that it can have an enormous size, and this may impede working with graphs possessing big size but simple structure. This drawback may be somehow rectified if another kind of matrix, "type adjacency matrix", is used. The type adjacency matrix often has a much smaller size in comparison with the adjacency matrix and reflects important features of the structure of the graph in an aggregate form. We will show that in many respects the type adjacency matrix may serve as a substitute for the adjacency matrix.

The presence of the type adjacency matrix corresponds to the fact that the set of nodes of undirected graph can be divided into a minimal number of disjoint subsets (types) in such way that each node of definite type has definite numbers of neighbors (adjacent nodes) of each type.

We provide definitions of the types of nodes and the type adjacency matrix and study the properties of the latter and its applications to several centrality measures. Also we introduce an analogue of the type adjacency matrix for directed weighted network and demonstrate its use in analysis of dynamic patterns in an economic development model based on the structure of optimal paths in directed networks.

2 Types of nodes and type adjacency matrix

2.1 Basic concepts

Let G be undirected connected graph with n nodes and $\mathbf{A}$ be its adjacency matrix, i.e. a $n \times n$ matrix such that $a_{ij} = a_{ji} = 1$ if in G there is an edge connecting nodes i and j, and $a_{ij} = a_{ji} = 0$ otherwise; $a_{ii} = 0$ for all $i = 1, 2, \ldots, n$.

The set of nodes, $N = \{1, 2, \ldots, n\}$ may be decomposed into minimal number S of disjoint classes N_i $(i = 1, 2, \ldots, S)$ in such way that any node belonging class i has t_{ij} neighbors from class j $(j = 1, 2, \ldots, S)$. The classes will be referred as *types of nodes*. Type i is characterized by vector $\mathbf{t}_i = (t_{i1}, t_{i2}, \ldots, t_{iS}$. We will consider a $S \times S$-matrix $\mathbf{T}$ with rows $\mathbf{t}_1, \ldots, \mathbf{t}_S$, which will be referred as *type adjacency matrix* of graph G.

Example 1. The adjacency matrix

$$\mathbf{A} = \begin{pmatrix} 0\,1\,1\,1\,1 \\ 1\,0\,1\,1\,0 \\ 1\,1\,0\,0\,1 \\ 1\,1\,0\,0\,1 \\ 1\,0\,1\,1\,0 \end{pmatrix}$$

describes a graph which has two types: $N_1 = \{1\}$ and $N_2 = \{2, 3, 4, 5\}$. The types are characterized by the numbers of neighbors $t_{11} = 0, t_{12} = 4, t_{21} = 1, t_{22} = 2$. The corresponding type adjacency matrix is

$$\mathbf{T} = \begin{pmatrix} 0\,4 \\ 1\,2 \end{pmatrix}.$$

Evidently, the familiar class of regular (equidegree) graphs is the class of one-type graphs. The next in the order of the generalized regular graphs are the classes of 2-types graphs, 3-type graphs, etc.

It is easy to see that there exist graphs of different size but with the same typology, and that for any (maybe small) number of types there exist graphs of unlimited size possessing given typology.

In any 2-types graph, the types correspond degrees of nodes. However, as the following example demonstrates, the distribution of degrees does not define typology.

Example 2. Consider two graphs with adjacency matrices

$$\mathbf{A_1} = \begin{pmatrix} 0\,1\,0\,0\,1\,0\,0\,1 \\ 1\,0\,1\,0\,0\,0\,0\,0 \\ 0\,1\,0\,1\,0\,0\,1\,0 \\ 0\,0\,1\,0\,1\,0\,0\,0 \\ 1\,0\,0\,1\,0\,1\,0\,0 \\ 0\,0\,0\,0\,1\,0\,1\,0 \\ 0\,0\,1\,0\,0\,1\,0\,1 \\ 1\,0\,1\,0\,0\,0\,1\,0 \end{pmatrix}, \mathbf{A_2} = \begin{pmatrix} 0\,1\,0\,0\,0\,0\,0\,1 \\ 1\,0\,1\,0\,0\,0\,0\,0 \\ 0\,1\,0\,1\,0\,0\,0\,1 \\ 0\,0\,1\,0\,1\,0\,1\,0 \\ 0\,0\,0\,1\,0\,1\,0\,0 \\ 0\,0\,0\,0\,1\,0\,1\,0 \\ 0\,0\,0\,1\,0\,1\,0\,1 \\ 1\,0\,1\,0\,0\,0\,1\,0 \end{pmatrix}.$$

Each of these graphs has 4 nodes of degree 2 and 4 nodes of degree 3. However, their typologies are different. The corresponding type adjacency matrices are

$$\mathbf{T_1} = \begin{pmatrix} 0 & 2 \\ 2 & 1 \end{pmatrix}, \mathbf{T_2} = \begin{pmatrix} 1 & 1 \\ 2 & 1 \end{pmatrix}.$$

2.2 Relation to other work

Our type adjacency matrix $\mathbf{T}$ is kin to a quotient matrix used in spectral graph theory (e.g. [6, 2]). Our subdivision of the set of nodes of graph into types is a generalization of the modular aggregation of Allouch [1], where *module* is defined as a subset of nodes, such that each node in a module has the same neighbors. It can be seen that each division into modules is a division into types, but not the opposite. Estrada [7] (section 2.4) provided examples of extension of the adjacency relations to links, cliques and other objects; however, his approach is rather far from our theory.

A polynomial-time algorithm of subdivision of the set of nodes into types is proposed by Matveenko and Korolev [11].

3 Some applications in network analysis

3.1 Numbers of walks

It is known that each element a_{ij}^k of the k-th power of the adjacency matrix, $\mathbf{A}^k$, is equal to the number of all k-step walks between nodes i and j. In a parallel way, each element t_{ij}^k of the k-th power of the type adjacency matrix, $\mathbf{T}^k$, is the number of all k-step walks between any i-th type node and all j-th type nodes.

3.2 Centrality measures

The *degree centrality* may be calculated in similar ways by use of either the adjacency matrix, $\mathbf{A}$, or the type adjacency matrix $\mathbf{T}$. The degrees of node i and of any node of type $\tilde{i}$ are

$$d_i = \sum_{j=1}^{n} a_{ij}, i = 1, 2, \ldots, n,$$

$$d = \sum_{j=1}^{S} t_{\tilde{i}j}, \tilde{i} = 1, 2, \ldots S.$$

If $i \in N_{\tilde{i}}$, then $d_i = d_{\tilde{i}}$.

Remind that the vector of *Bonacich centrality measures* for nodes [4] is

$$\mathbf{C}^B = (\mathbf{I} - \alpha\mathbf{A})^{-1} \cdot \mathbf{1},$$

where $\mathbf{I}$ is the $n \times n$ unit matrix, $\mathbf{1}$ is the n-vector of all ones; $0 M\alpha < 1$. The vector

$$\mathbf{C}^K = ((\mathbf{I} - \alpha\mathbf{A})^{-1} - \mathbf{I}) \cdot \mathbf{1},$$

is often referred as vector of *Katz centrality measures* [8, 3].

Since only the summary discounted numbers of walks from node i (but not of walks to particular nodes) are needed in calculation of both Katz-Bonacich centrality measures, the type adjacency matrix, $\mathbf{T}$, can be used instead of the adjacency matrix, $\mathbf{A}$. We obtain the vector of Bonacich centrality measures for types:

$$\tilde{\mathbf{C}}^B = (\tilde{\mathbf{I}} - \alpha\mathbf{T})^{-1} \cdot \tilde{\mathbf{1}},$$

where $\tilde{\mathbf{I}}, \mathbf{T}$ are $S \times S$-matrices, and $\tilde{\mathbf{1}}$ is the S-vector of all ones; $0 < \alpha < 1$. Correspondingly, the vector of Katz centrality measures for types is

$$\tilde{\mathbf{C}}^K = ((\tilde{\mathbf{I}} - \alpha\mathbf{T})^{-1} - \tilde{\mathbf{I}})\tilde{\mathbf{1}}.$$

It is easy to show that the vector of α-centralities [5],

$$\mathbf{k} = (\mathbf{I} + \alpha\mathbf{A})^{-1}\mathbf{e},$$

where α is any number and e is any vector, can be calculated for types by use of the type adjacency matrix, $\mathbf{T}$:

$$\tilde{\mathbf{k}} = (\tilde{\mathbf{I}} + \alpha\mathbf{T})^{-1}\tilde{\mathbf{e}}.$$

Here $\tilde{\mathbf{k}}$ is the S-vector of α centralities for types, $\tilde{\mathbf{I}}$ is the $S \times S$ unit matrix, $\tilde{\mathbf{e}}$ is S-vector such that $\tilde{e}_{\tilde{i}} = e_i$ if $i \in N_{\tilde{i}}$.

The following theorem implies that the type adjacency matrix $\mathbf{T}$ can be used instead of the adjacency matrix $\mathbf{A}$ for calculation of *eigenvalue centralities*.

Theorem 1. *1. Let λ be an eigenvalue of the type adjacency matrix $\mathbf{T}$ and $\tilde{\mathbf{b}}$ be corresponding eigenvector. Then λ is also an eigenvalue of the adjacency matrix $\mathbf{A}$, and the corresponding eigenvector is $\mathbf{b}$, where $b_i = \tilde{b}_{\tilde{i}}$ if $i \in N_{\tilde{i}}$.*

2. Let λ_F be the Frobenius eigenvalue of matrix $\mathbf{T}$. Then λ_F is also the Frobenius eigenvalue of matrix $\mathbf{A}$.

Proof. Part 1 follows directly from the definitions. To prove ad absurdum part 2, assume that the Frobenius eigenvalue of matrix $\mathbf{A}$ is $\mu \neq \lambda_F$. Part 1 implies $\mu > \lambda_F$. Let $\mathbf{e}$ be the Frobenius eigenvector of matrix $\mathbf{A}$, and $\hat{\mathbf{e}}$ be the n-vector with components

$$\hat{e}_i = \max_{j \in N_{\tilde{i}}} e_j,$$

where $i \in N_{\tilde{i}}$. Evidently,

$$\mathbf{A}\hat{\mathbf{e}} \geq \mu\hat{\mathbf{e}}.$$

Let $\hat{\mathbf{f}}$ be the S-vector corresponding to $\hat{\mathbf{e}}$ (i.e. $\hat{f}_{\tilde{i}} = \hat{e}_i$ if $i \in N_{\tilde{i}}$). Then

$$\mathbf{T}\hat{\mathbf{f}} \geq \mu\hat{\mathbf{f}}. \tag{1}$$

But, according to the Perron-Frobenius theorem, since λ_F is the Frobenius eigenvalue, (1) implies $\lambda_F \geq \mu$. Contradiction! $\qquad\square$

3.3 Coincidence of the orders generated by different centrality measures

An open question in network analysis is existence of graph classes for which different centrality measures do order the nodes in the same way. One of such classes (trees with "monotone hierarchies") is found by Bloch et al. [3]. We find that our class of 2-types graphs is another such class.

Notice that it can be shown that the Bonacich centrality exists iff $0 < \alpha < \lambda_F$. We assume that this condition is satisfied.

Theorem 2. *For nodes $i = 1, 2, \ldots, n$ let d_i, C_i^e, C_i^B be, correspondingly, degree, the eigenvalue centrality and the Bonacich centrality. For any 2-types graph, the following statements for a pair of nodes i, j are equivalent:*

1) $d_i > d_j$,
2) $C_i^e > C_j^e$,
3) $C_i^B > C_j^B$.

Proof. 1-2). Vector $(1, x)^T$, where $x = C_2^e/C_1^e$, is the Frobenius eigenvector of matrix $\mathbf{T}$; hence, $t_{11} + t_{12}x = \lambda_F$. Thus, $x > 1$ iff $\lambda_F > d_1$. On the other hand, $\lambda_F \in (\min\{d_1, d_2\}, \max\{d_1, d_2\})$, which implies equivalence of inequalities $\lambda_F > d_1$ and $d_2 > d_1$. Hence, $C_2^e > C_1^e$ iff $d_2 > d_1$.

1-3). The definition of the Bonacich centrality implies that the vector of centralities for types is

$$C^B = \frac{1}{\Delta} \left(1 + \alpha(t_{12} - t_{22}), 1 + \alpha(t_{21} - t_{11}) \right)^T,$$

where $\Delta = 1 - \alpha Tr\mathbf{T} + \alpha^2 Det\mathbf{T}$. It follows that

$$C_1^B - C_2^B = \frac{1}{\Delta}(d_1 - d_2).$$

This implies

$$(C_1^B - C_2^B)\alpha \left(\frac{1}{\alpha} - \lambda_F \right) \left(\frac{1}{\alpha} - \lambda_2 \right) = d_1 - d_2,$$

where λ_2 is the smallest eigenvalue of matrix $\mathbf{T}$. The existence condition implies

$$\frac{1}{\alpha} > \lambda_F > \lambda_2.$$

Hence, the differences $C_1^B - C_2^B$ and $d_1 - d_2$, have the same sign. □

The list of the centrality measures in Theorem 2 may be enlarged, but in this short paper we limit ourselves only by the three measures.

4   Typology of directed graphs

Let G be a strongly connected directed weighted graph with n nodes, and $\mathbf{A}$ be its $n \times n$ weights matrix. If there is no arc going from node i to node j, then it is taken $a_{ij} = +\infty$. Types $1, 2, \ldots S$ can be defined as components of disjoint division of the set of nodes $\{1,2,\ldots, n\}$ such that there is a $S \times S$-matrix of weights, $\mathbf{T}$, where $t_{\tilde{i}\tilde{j}} = a_{ij}$ if nodes i, j are of types $\tilde{i}, \tilde{j}$ correspondingly. Matrix $\mathbf{T}$, referred further as *type weights matrix*, defines a *types graph*, H, with S nodes.

4.1   Application to a model of economic development

Let G be a network of economic agents (e.g. firms, authorities, groups of workers of different qualification types, etc.) Time is discrete, $t = 0, 1, \ldots$ Each agent i $(i = 1, 2, \ldots, n)$ at period t is characterized by a positive number x_i^t referred as *value* (e.g. profit, income, welfare). The agent's current value depends on the values of herself and her neighbors in the previous time period. Following [10], we assume that the development of each agent i $(i = 1, 2, \ldots)$ is limited:

1) by her own potential possibilities described by coefficient $a_{ij} > 0$:

$$x_i^{t+1} \le a_{ii} x_i^t, t = 0, 1, \ldots;$$

2) by the sizes of externalities created by her neighbors, described by coefficients $a_{ij} > 0$:

$$x_i^{t+1} \le a_{ij} x_j^t, j = 1, 2, \ldots, n; j \neq i; t = 0, 1, \ldots.$$

We build matrix $\mathbf{A}$ of the coefficients, assuming that $a_{ij} = +\infty$ if there is no arc going from node i to node j, or if agent j does not limit anyway the agent i's development.

Each agent maximizes her value; this leads to the following equations characterizing the equilibrium path given initial values x_i^0:

$$x_i^{t+1} = \min_{j=1,2,\ldots,n} a_{ij} x_j^t, i = 1, 2, \ldots, n$$

The dynamic system can be written as

$$\mathbf{x}^{t+1} = \mathbf{A}\mathbf{x}^t = \mathbf{A}^t \mathbf{x}^0, \tag{2}$$

where $\mathbf{x}^t$ is n-vector with positive components x_i^t, and the matrix-vector multiplication and the matrix power are in sense of tropical/idempotent mathematics with operation $+ = \min$ and usual multiplication.

The dynamic system (2) may demonstrate various patterns (e.g. entrance to a steady state, to a ray of stable growth or decline, to a limit cycle, etc). The pattern is totally defined by the structure of matrix $\mathbf{A}$ and, correspondingly, of network G (see [9, 10]) .

Let us consider the set Σ of (oriented) cycles in graph G. To each cycle $\sigma = (i_1, i_2, \ldots, i_k)$ an *average weight* corresponds: $a_\sigma = (a_{i_1 i_2} a_{i_2 i_3} \ldots a_{i_k i_1})^{1/k}$. A cycle σ^* is called *optimal* if

$$a_{\sigma^*} = a^* = \min_{\sigma \in \Sigma} a_\sigma.$$

In a similar way we can define optimal cycle in the types graph H and the corresponding minimal average weight, t^*.

Theorem 3. $a^* = t^*$.

Proof. Evidently, to each cycle in graph G a unique cycle in the types graph H corresponds. Vice versa, to each cycle in the types graph H at least one cycle in graph G corresponds. It is easily checked that the average weights of these cycles do coincide. Hence, the minimal average weights also do coincide. □

It is shown in [9, 10] that a^* is the long-run growth factor of the dynamic system (2). Hence, by Theorem 3, the growth factor is defined by the type weights matrix $\mathbf{T}$.

Example 3. In case of 2×2 type weights matrix $\mathbf{T}$, the long-run growth factor is

$$a^* = t^* = \min\{t_{11}, t_{22}, \sqrt{t_{12} t_{21}}\}.$$

If $t^* > (<)1$, there is a long-run growth (decline) in system (2).

References

1. Allouch, N.: Aggregation in networks. Queen Mary University of London. School of Economics and Finance. (2016)
2. Atik, F., Panigrahi, P.: Graphs with few distinct distance eigenvalues irrespective of the diameters. Electr. J. Lin. Alg. 29, 194–205 (2015)
3. Bloch, F., Jackson, M.O., Tebaldi, P.: Centrality measures in networks. arXiv: 1608.05845 (2017)
4. Bonacich, P.: Factoring and weighting approaches to status scores and clique identification. J. of Math. Sociology 2, 113–120 (1972)
5. Bonacich, P., Lloyd, P.: Eigenvector-like measures of centrality for asymmetric relations. Soc. Networks 23, 191–201 (2001)
6. Brouwer, A.E., Haemers W.H.: Spectra of graphs. Springer, New York (2012)
7. Estrada, E. The structure of complex networks. Theory and applications. Oxford University Press, Oxford (2011)
8. Goyal, S.: Connections: an introduction to the economics of networks. Princeton University Press, Princeton (2009)
9. Matveenko, V.: Optimal trajectories of a dynamic programming scheme and extremal degrees of nonnegative matrices (In Russian). Diskretnaya Matematika 2, 59–71 (1990)

10. Matveenko, V.: Development with positive externalities: the case of the Russian economy. J. Pol. Modeling 17, 207–221 (1995)
11. Matveenko, V., Korolev, A.: Equilibria in networks with production and knowledge externalities. Springer Proc. Math. Stat. 156, 291–331 (2016)