Construction of crisis precursors in multiplex networks

Vladimir Soloviev
Department of Computer Science and Applied Mathematics
State Pedagogical University
Kryvyi Rih, Ukraine
vnsoloviev2016@gmail.com

Viktoria Solovieva*
Department of Management and Public Administration
Kryvyi Rih Economic Institute of Kyiv
National Economic University named after Vadim Hetman
Kryvyi Rih, Ukraine
vikasolovieva2027@gmail.com

Anna Tuliakova
Department of Higher Mathematics
Odessa National University named after I. I. Mechnikov
Odessa, Ukraine
TuliakovaAnna@gmail.com

Margaryta Ivanova
Department of Entrepreneurship and Marketing
Ivano-Frankivsk National Technical University Oil and Gas
Ivano-Frankivsk, Ukraine
margaretivanova@gmail.com

Abstract—Based on the network paradigm of complexity in the work, a systematic analysis of the dynamics of the largest stock markets in the world has been carried out. According to the algorithms of the visibility graph and recurrence plot, the daily values of stock indices are converted into a multiplex networks, the spectral and topological properties of which are sensitive to the critical and crisis phenomena of the studied complex systems. It is shown that some of the spectral and topological characteristics can serve as measures of the complexity of the stock market, and their specific behaviour in the pre-crisis period is used as indicators-precursors of crisis phenomena.

Keywords—stock markets, graph theory, complex networks.

I. INTRODUCTION

The new interdisciplinary study of complex systems, known as the complex networks theory, laid the foundation for a new network paradigm of synergetics [1]. The complex networks theory studies the characteristics of networks, taking into account not only their topology, but also statistical properties, the distribution of weights of individual nodes and edges, the effects of information dissemination, robustness, etc. [2-8]. Complex networks include electrical, transport, information, social, economic, biological, neural and other networks [3-8]. The network paradigm has become dominant in the study of complex systems since it allows you to enter new quantitative measures of complexity not existing for the time series [9].

Previously, we introduced various quantitative measures of complexity for individual time series [10, 11]. Significant advantage of the introduced measures is their dynamism, that is, the ability to monitor the time of change in the chosen measure and compare with the corresponding dynamics of the output time series. This allowed us to compare the critical changes in the dynamics of the system, which is described by the time series, with the characteristic changes of concrete measures of complexity. It turned out that quantitative measures of complexity respond to critical changes in the dynamics of a complex system, which allows them to be used in the diagnostic process and prediction of future changes. In [11], we introduced network complexity measures and adapted them to study system dynamics. But networks are rarely isolated. Therefore, it is necessary to take into account the interconnection interaction, which can be realized within the framework of different models [12]. In this paper we will consider it by simulating so-called multiplex networks, the features of which are reduced to a fixed number of nodes in each layer, but they are linked by different bonds.

II. METHODS OF CONVERTING TIME SERIES INTO GRAPHS

Most complex systems inform their structural and dynamic nature by generating a sequence of certain characteristics known as time series. In recent years, interesting algorithms for the transformation of time series into a network have been developed, which allows to extend the range of known characteristics of time series even to network ones. Recently, several approaches have been proposed to transform time sequences into complex network-like mappings. Three main classes can be distinguished. The first is based on the study of the convexity of successive values of the time series and is called visibility graph (VG) [13]. The second analyzes the mutual approximation of different segments of the time sequence and uses the technique of recurrent analysis [14]. The recurrent diagram reflects the existing repetition of phase trajectories in the form of a binary matrix whose elements are units or zeros, depending on whether they are close (recurrent) with given accuracy or not, the selected points of the phase space of the dynamic system. The recurrence diagram is easily transformed into adjacency matrix, on which the spectral and topological characteristics of the graph are calculated [14]. Finally, if the basis of forming the links of the elements of the graph is to put correlation relations between them, we obtain a correlation graph [14]. To construct and analyze the properties of a correlation graph, we must form adjacency matrix from the correlation matrix. To do this, you need to enter a value which, for the correlation field, will serve as the
distance between the correlated agents. So, if the correlation coefficient between the two assets is significant, the distance between them is small, and, starting from a certain critical value, assets can be considered bound on the graph. For an adjacency matrix, this means that they are adjacent to the graph. Otherwise, the assets are not contiguous. In this case, the binding condition of the graph is a prerequisite.

The use of the complexity of recurrent and visibility graph networks to prevent critical and crisis phenomena in stock markets has been considered by us in a recent papers [15, 16]. In this paper, we consider multiplex implementations of these techniques.

The recurrence diagrams for the visualization of phase space recurrences is based on Henri Poincare's idea of the phase space recurrence of dynamical systems. According to Takens theorem [14], an equivalent phase trajectory that preserves the structure of the original phase trajectory can be recovered from a single observation or time series by the time delay method: \( x(t) = \langle u_1, u_2, ..., u_{m+1} \rangle \), where \( m \) - embedding dimension, \( \tau \) - time delay (real time delay is defined as \( \tau \Delta t \)). The recurrent diagram shows the existing repetitions in the form of a binary matrix \( R \), where \( R_{ij} = 1 \), if \( \bar{x}_i \) it is adjacent to the state \( \bar{x}_j \), and \( R_{ij} = 0 \) otherwise. Neighboring (or recurrent) are states \( \bar{x}_i \) that fall into a \( m \) - dimensional region with radius \( \varepsilon \) and center in \( \bar{x}_j \). It is clear that parameters \( m \), \( \tau \) and \( \varepsilon \) are key when conducting recurrent analysis. The recurrence diagram is easily transformed into an adjacency matrix, by which the spectral and topological characteristics of the graph are calculated [15].

The algorithm of the VG is realized as follows. Take a time series \( Y(t) = \{y_1, y_2, ..., y_N\} \) of length \( N \). Each point in the time series data can be considered as a vertex in an associative network, and the edge connects two vertices if two corresponding data points can "see" each other from the corresponding point of the time series (Figure 1).

Formally, two values \( y_a \) of the series (at the time of time \( t_a \)) and \( y_b \) (at the time of time \( t_b \)) are connected, if for any other value \( y_c \), which is placed between them (that is, \( t_a < t_c < t_b \)), the condition is satisfied \( y_a < y_c + (y_c - y_b)(t_a - t_c)/(t_b - t_c) \). Note that the visibility graph is always connected by definition and also is invariant under affine transformations, due to the mapping method.

An alternative (and much simpler) algorithm is the horizontal visibility graph (HVG) [17], in which a connection can be established between two data points \( a \) and \( b \), if one can draw a horizontal line in the time series joining them that does not intersect any intermediate data by the following geometrical criterion: \( y_a > y_b > y_c \) for all \( c \) such that \( t_a < t_c < t_b \) (Figure 1).

In multiplex networks, there are two tasks [18]: (1) turn separate time series on the network for each layer; (2) connect the intra-loop networks to each other. The first problem is solved within the framework of the standard algorithms described above. For multiplex networks, the algorithm of the MVG for the three layers is presented in Figure 2.

The cross-recurrent multiplex network (MCRN) is formed from recursive diagrams of individual layers.

### III. SPECTRAL AND TOPOLOGICAL GRAPH PROPERTIES

Spectral theory of graphs is based on algebraic invariants of a graph - its spectra. The spectrum of graph \( G \) is the set of eigenvalues of a matrix \( \Sigma_2(G) \) corresponding to a given graph. For an adjacency matrix \( A \) of a graph, there exists a characteristic polynomial \( |\lambda I - A| \), which is called the characteristic polynomial of a graph \( P_\lambda(I) \). The eigenvalues of the matrix \( A \) (the zeros of the polynomial \( |\lambda I - A| \)) and the spectrum of the matrix \( A \) (the set of eigenvalues) are called respectively their eigenvalues and the spectrum of graph \( G \). The eigenvalues of the matrix \( A \) satisfy the equality \( \lambda \mathbf{x} = \lambda \mathbf{x} \) (\( \mathbf{x} \) - non-zero vector). Vectors \( \mathbf{x} \) satisfying this equality are called eigenvectors of matrix \( A \) (or graph \( G \)) corresponding to their eigenvalues.

Another common type of graph spectrum is the spectrum of the Laplace matrix \( L \). The Laplace matrix is used to calculate the tree graphs, as well as to obtain some important spectral characteristics of the graph. In particular, the positive eigenvalues \( \lambda_2 \) is called the index of algebraic connectivity of the graph. This value represents the "force" of the connectivity of the graph component and is used in the analysis of reliability and synchronization of the graph.

Important derivative characteristics are spectral gap, graph energy, spectral moments and spectral radius. The spectral gap is the difference between the largest and the next eigenvalues of the adjacency matrix and characterizes the rate of return of the system to the equilibrium state. The graph energy is the sum of the modules of the eigenvalues of the graph adjacency matrix. The spectral radius is the largest modulus of the eigenvalue of the adjacency matrix. Denote by \( N_e \) the value which corresponds to an "average eigen
where. For most networks there is a power law distribution \( P(k) \propto k^{-\gamma} \), where \( \gamma > 0 \).

Also important topological characteristics are the vertex eccentricity - the largest distance between \( m \) and any other vertex, that is, how far the vertex is far from the other vertices of the graph. The centrality of the vertex measures its relative importance in the graph. At the same time, the farness of a node is defined as the sum of its distances to all other nodes, and its closeness is defined as the backward distance. Thus, the centrality of the node is lower than its total distance to all other nodes.

Another important measure is the link density in the graph, which is defined as the number of links \( n_e \), divided by the expression \( n_e (n_e - 1) / 2 \), where \( n_e \) is the number of nodes of the graph.

IV. MULTIPLEX COMPLEXITY MEASURES

A multilayer/multiplex network is a pair \( M = (G, C) \) where \( G = \{G_{\alpha}; \alpha \in \{1, \ldots, M\}\} \) there is a family of graphs (whether directed or not, weighted or not) \( G_{\alpha} = \{X_{\alpha}, E_{\alpha}\} \), called layers \( C = \{E_{\alpha \beta} \subseteq X_{\alpha} \times X_{\beta} ; \alpha, \beta \in \{1, \ldots, M\}, \alpha \neq \beta \} \). The latter is a set of links between nodes of different layers \( G_{\alpha} \) and \( G_{\beta} \) at \( \alpha \neq \beta \). Each element \( E_{\alpha \beta} \) is an interlayer bond \( M \) in contrast to the elements of each \( E_{\alpha \beta} (\alpha \neq \beta) \), called interlayer bonds.

A set of nodes of a layer \( G_{\alpha} \) is denoted \( X_{\alpha} = \{x_1^\alpha, \ldots, x_{N_{\alpha}}^\alpha\} \), and a adjacency matrix \( A_{\alpha} \) as \( A_{\alpha}^{ij} = \left( a_{\alpha}^{ij} \right)_{N_{\alpha} \times N_{\alpha}} \), where \( a_{\alpha}^{ij} = \begin{cases} 1, & (x_i^\alpha, x_j^\alpha) \in E_{\alpha} \setminus E_{\beta} \setminus E_{\alpha \beta} \\ 0, & \text{otherwise} \end{cases} \) for \( 1 \leq i, j \leq N_{\alpha} \) and \( 1 \leq \alpha \leq M \). For an interlayer adjacency matrix we have \( A_{\alpha \beta}^{ij} = \left( a_{\alpha \beta}^{ij} \right)_{N_{\alpha} \times N_{\beta}} \), where \( a_{\alpha \beta}^{ij} = \begin{cases} 1, & (x_i^\alpha, x_j^\beta) \in E_{\alpha \beta} \\ 0, & \text{otherwise} \end{cases} \)

A multiplex network is a partial interlayer case and contains a fixed number of nodes connected by different types of links. Multiplex networks are characterized by correlations of different nature [16], which enable the introduction of additional multiplexes.

Let’s evaluate the quantitative overlap between the various layers. The average edge overlap obviously equal

\[
\omega = \frac{\sum_{i,j} \sum_{\alpha \beta} \sum_{a_{\alpha \beta}^{ij}} \frac{1}{M} \sum_{\alpha \beta} \sum_{a_{\alpha \beta}^{ij}} (1 - \delta_{\alpha \beta} \sum_{a_{\alpha \beta}^{ij}})}{1/M \text{ if the connection } (i, j) \text{ exists only in one layer, that is, if there is a layer } \alpha \text{ such that } a_{\alpha \beta}^{ij} = 1, \text{ } a_{\alpha \beta}^{ij} = 0 \text{ } \forall \beta \neq \alpha .
\]

If all layers are identical, then \( \omega = 1 \). Consequently, this measure can serve as a measure of the coherence of the
output time series: high values \( \omega \) indicate a noticeable correlation in the structure of time series.

The total overlap \( O^{[\alpha]} \) between the two layers \( \alpha \) and \( \beta \) is defined as the total number of bonds that are shared between the layers \( \alpha \) and \( \beta \): \( O^{[\alpha]} = \sum_{i} a_{i}^{[\alpha]} a_{i}^{[\beta]} \), where \( \alpha \neq \beta \).

For a multiplex network, the vertex degree \( k \) is already a vector \( k = (k^{[1]}, \ldots, k^{[M]}) \), where \( k^{[\alpha]} \) is the degree of the node in the layer, that is, while the elements of the matrix of adjacency for the layer. Specificity of the vector character of the degree of the peak in multiplex networks allows for the introduction of additional interlayer characteristics. One of these is the overlap of the node's degree \( O_{i} = \sum_{\alpha} k_{i}^{[\alpha]} \).

The next measure quantitatively describes the interlayer correlations between the degrees of the selected node in two different layers. If, chosen from \( M \) the layers of the pair \((\alpha, \beta)\) characterized by the distribution of degrees \( P(k_{i}^{[\alpha]}), P(k_{i}^{[\beta]}) \), the so-called interlayer mutual information is determined by the formula

\[
I_{\alpha, \beta} = \sum_{i} P(k_{i}^{[\alpha]}, k_{i}^{[\beta]}) \log \left( \frac{P(k_{i}^{[\alpha]}k_{i}^{[\beta]})}{P(k_{i}^{[\alpha]}P(k_{i}^{[\beta]}))} \right),
\]

where \( P(k_{i}^{[\alpha]}, k_{i}^{[\beta]}) \) is the probability of finding a node degree \( k_{i}^{[\alpha]} \) in a layer \( \alpha \) and degree \( k_{i}^{[\beta]} \) in a layer \( \beta \). The higher the \( I_{\alpha, \beta} \) value, the more correlated are the distributions of the levels of the two layers, and, consequently, the structure of the time series associated with them. We also find the mean value \( \bar{I}_{\alpha, \beta} \) for all possible pairs of layers - the scalar value \( \langle I_{\alpha, \beta} \rangle \) that quantifies the information flow in the system.

The quantity that quantitatively describes the distribution of the node degree \( j \) between different layers is the entropy of the multiplexed degree \( S_{j} = \sum_{\alpha} k_{i}^{[\alpha]} \log \left( k_{i}^{[\alpha]} / O_{\alpha} \right) \).

Entropy is zero if all nodes are in the same layer and vice versa, has the maximum value when they are evenly distributed between different layers. That is, the higher the value \( S_{j} \), the even more links evenly the nodes' connections are distributed between the layers.

A similar magnitude is the multiplex participation coefficient

\[
P_{i} = M / (M - 1) \left[ 1 - \sum_{\alpha} \left( k_{i}^{[\alpha]} / O_{\alpha} \right) \right].
\]

\( P_{i} \) takes values on the interval \([0,1]\) and determines that homogeneous links of node \( i \) are distributed among \( M \) the layers. If all links of the node \( i \) lie in one layer, \( P_{i} = 0 \) and \( P_{i} = 1 \) if the node has a precisely defined number of links in each of the \( M \) layers. Consequently, the larger the coefficient \( P_{i} \) is, the more evenly distributed the participation of the node in the multiplex.

Obviously, the magnitudes \( S_{j} \) and \( P_{i} \) are very similar. We will show that some of these spectral and topological measures serve as measures of complexity of the system, and the dynamics of their changes allows us to build predictors of crisis situations on financial markets.

V. EXPERIMENTAL RESULTS AND THEIR DISCUSSION

The time series of daily values of stock market indexes for the period from 01/01/1983 to 10/01/2019 were selected as databases, which contained significant changes in the indexes, and were identified as crisis phenomena [23]. Among the set of stock indexes are the following:

- SP (S & P 500) - USA;
- DAX (DAX PERFORMANCE-INDEX) - Germany;
- N225 (Nikkei 225) - Japan;
- HSI (HANG SENG INDEX) - China;
- GSPTSE (S & P / TSX Composite index) - Canada.

The size of the databases was determined by the shortest available database [23], which in this case was determined by the length of the index DAX (Figure 3).

![Fig. 3. The dynamics of the daily values of the selected stock market indices from December 4, 1987 to August 22, 2019.](image)

Despite a slight difference in the dynamic stock indexes (this is especially noticeable for the index NIKKEY), the main trends (both rising and falling – actually crises) are observed for all time series.

Calculations of spectral and topological measures by methods of MVG, MCRP were carried out in the following way. The time window was chosen, for example, a year or two (approximately 250 or 500 trading days), for which the corresponding graphs were constructed and the spectral, topological and multiplex properties were calculated. Next, the window was shifted step by step, for example, one week (5 trading days) and the procedure repeated until the time series were exhausted. Knowing the time of the onset of the crisis and comparing the time series with the dynamics of a certain indicator, it is possible to investigate its dependence on certain the characteristic changes in the stock market: pre-crisis, crisis and post-crisis periods.

The results of calculations for revived time series of graphs are shown in Figures 4 - 7.
Figures 4 - 5 show that the pair of parameters $S, P$ in Figure 4 is antisymmetric to the three parameters $I, o, O$ of Figure 5. However, all of them in a characteristic way (decreasing or increasing before the crisis) signal about its possible onset.

As in the case of a multiplex visibility graph, multiplex recurrence measures are also informative indicators-predictors of crisis phenomena.

Parameters such as the width of the window $w$ and the step $\Delta t$ of its displacement along the time series are important. When $w$ is small, the degree of complexity fluctuates noticeably, reacting not only to crises, but also to more or less noticeable fluctuations of the index. On the contrary, with too much window width there is a noticeable smoothing of the appropriate measure and if two crises are at a distance that is smaller $w$, the indicators of both crises are averaged and less informative. If you choose an oversized parameter $\Delta t$, you might miss the actual crisis that distorts the indicator.

Multiplex networks are actively used to simulate complex networks of different nature: from financial (banks [19], stock market [20], guarantee market [21]) to social [22]. Particular attention should be paid to the work [20], in which the above multiplex measures are analyzed for the subject of correlations with known stock markets crises. Yet there is no systematic analysis of network and multiplex measures and the construction of indicators-predictors of the crisis phenomena in the stock market. Therefore, our studies to some extent fill this gap.

VI. CONCLUSIONS

Thus, we have demonstrated the possibility of studying complex socio-economic systems as part of a network paradigm of complexity. A time series can be represented in an equivalent way - a multiplex network, which has a wide range of characteristics; both spectral and topological, and multiplexed. Examples of known financial crises have shown that most of the network measures can serve as indicators-precurors of crisis phenomena and can be used for possible early prevention of unwanted crises in the financial markets. They are an extension of the already proposed by us and "working" indicators, which use other measures of complexity.

It should be noted that the proposed indicators-precurors do not solve the more general problem of forecasting future values or trends of the stock market. In this way, it is possible to use new approaches or alternative methods based on algorithms of (deep) machine learning [24].
REFERENCES


