

# A General Approach to Network Analysis of Statistical Data Sets

Valery A. Kalygin<sup>1</sup>(✉), Alexander P. Koldanov<sup>1</sup>, and Panos M. Pardalos<sup>2</sup>

<sup>1</sup> Higher School of Economics, National Research University,  
Nizhny Novgorod, Russia  
{vkalyagin, akoldanov}@hse.ru

<sup>2</sup> Center for Applied Optimization, University of Florida, Gainesville, FL, USA  
pardalos@ufl.edu

**Abstract.** The main goal of the present paper is the development of general approach to network analysis of statistical data sets. First a general method of market network construction is proposed on the base of idea of measures of association. It is noted that many existing network models can be obtained as a particular case of this method. Next it is shown that statistical multiple decision theory is an appropriate theoretical basis for market network analysis of statistical data sets. Finally conditional risk for multiple decision statistical procedures is introduced as a natural measure of quality in market network analysis. Some illustrative examples are given.

**Keywords:** Market network analysis · Statistical data sets · Measures of association · Multiple decision theory

## 1 Introduction

Network analysis is a popular and powerful tool of modern analysis of complex systems [14, 15]. This analysis is known to be very useful for technological, social, biological, and other complex system. Nodes (vertices) of the network correspond to the elements of the complex system and links (edges) of the network correspond to the interaction between elements. Measure of interaction between nodes gives the weights of the links. Resulting weighted graph represents the network model of the complex system. The structure of the network is defined by the data sets that we use to measure the links. In the present paper we consider network models generated by statistical data sets. Important examples are market networks and brain connectivity networks. The statistical origin of the data generates error in the decision about network structures. This error can leads to erroneous interpretation of network analysis. The majority of existing publications in the field in our knowledge does not pay attention to this problem.

---

The authors are partly supported by National Research University Higher School of Economics, Russian Federation Government grant, N. 11.G34.31.0057 and RFFI 14-01-00807.

The main goal of the present paper is to develop a general approach to network analysis of statistical data sets in order to handle the related statistical errors.

Financial market is known to be a complex system. The complexity of the system is reflected in the associated complete weighted graph. The minimum spanning tree (MST) of the graph was studied in [13] to extract the most valuable information from this complex network. This information can be extended with the use of planar maximally filtered graph (PMFG) as suggested in [23]. Both procedures (MST and PMFG) can be considered as a filtering of a complex graph into a simpler relevant subgraph. Research in this direction is very active our days (see for example [24] where the state of art is given). Another filtering procedure was proposed in [3]. As a result of this procedure a *market graph* (MG) is constructed. Maximum cliques (MC) and maximum independent sets (MIS) of the market graph give an interesting information about financial market structures [4, 5] (for calculation of MC and MIS see [17, 18]).

The financial market has a large element of randomness. The scientific approach to handle the randomness of the financial market consists among others of the following connected stages:

- Design of the model of the market network, choice of the filtered structural characteristic (FSC).
- Identification of FSC from the observations, construction of appropriate statistical procedures.
- Control of uncertainty of statistical procedures.

It is common knowledge that the prices and returns of stocks of financial market are modeled by stochastic process [21]. A complete information about this process is given by the associated probabilistic space  $(\Omega, \mathfrak{F}, P)$ . It follows from the Kolmogorov consistency theorem that the process is defined by the collections of finite-dimensional joint distributions. To model the associated network one has to introduce a measure of interaction between stocks. Any measure of interaction (dependence) between stocks therefore has to be extracted from the joint distributions. This give rises to the concept of *true market network and true FSC*. Once the measure of interaction is defined one can go to the next stage: identification of the market network and FSC from observations. This gives rise to the concept of *sample market network and sample FSC*. Control of uncertainty can be based now on the analysis of the difference between true market network and sample market network and true FSC and sample FSC.

In the present paper we develop a general approach which generalizes some ideas from [1, 2, 6–9]. First we propose a general approach to design a different models for market network on the base of idea of *measure of association* introduced in [10] and developed in [11]. We show that existing network models [13, 16, 19] can be obtained from this approach. Next we show that statistical multiple decision theory is an appropriate theoretical basis for identification of filtered structural characteristic (FSC). Finally we introduce the conditional risk as a natural measure of quality in market network analysis.

The paper is organized as follows. In Sect. 2 we describe some class of measures of dependence that we call measures of association. In Sect. 3 we discuss

identification problem for filtered structural characteristics (FSC). In Sect. 4 we put the market network analysis in the framework of multiple decision theory. In Sect. 5 we discuss the conditional risk as a measure of quality in market network analysis and give some illustrative examples.

## 2 Measures of Association

There are many measures of dependence between two random variables proposed in the literature: Pearson correlation, Kruskal correlation, Kendall correlation, Spearman correlation, Fehner correlation and others [22]. Many of them can be put in the framework of the general concept proposed in [11]. According to Lehmann, random variables  $X, Y$  are positively dependent if

$$P(X \leq x, Y \leq y) \geq P(X \leq x)P(Y \leq y), \text{ for all } (x, y) \in R^2 \quad (1)$$

In terms of the joint distribution function this reads

$$F_{X,Y}(x, y) - F_X(x)F_Y(y) \geq 0, \text{ for all } (x, y) \in R^2 \quad (2)$$

Similarly,  $X, Y$  are negatively dependent if (1), (2) holds with inequality sign reversed. The definition of positive dependence compares the probability of the product of events with the product of probabilities of events in the sense that small value of  $Y$  tends to be associated with small value of  $X$  and (see below) large value of  $Y$  with large value of  $X$ . Dependence measures based on this comparison will be called in this paper *measures of association*. In particular covariance between two random variables is a measure of association as it follows from the Hoeffding formula [11]:

$$\text{Cov}(X, Y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [F_{X,Y}(x, y) - F_X(x)F_Y(y)] dx dy \quad (3)$$

It implies that if two random variables are positively dependent then their covariance and therefore Pearson correlation between them is non negative. Converse is known to be true for the normal vector  $(X, Y)$  [11]. It means that for the normal case positiveness of the correlation implies the positive dependence of the random variables. It gives a strong additional justification for the use of Pearson correlation as a measure of dependence in the normal case.

The condition (1) is equivalent to any of the following conditions

$$P(X \leq x, Y \geq y) \leq P(X \leq x)P(Y \geq y), \text{ for all } (x, y) \in R^2$$

$$P(X \geq x, Y \leq y) \leq P(X \geq x)P(Y \leq y), \text{ for all } (x, y) \in R^2$$

$$P(X \geq x, Y \geq y) \geq P(X \geq x)P(Y \geq y), \text{ for all } (x, y) \in R^2$$

Therefore if two variables  $X, Y$  are positively dependent then for any  $x, y \in R$  one has

$$P((X - x)(Y - y) > 0) - P((X - x)(Y - y) < 0) \geq 0$$

This observation produces a family of different measures of association  $q(x, y)$ :

$$q_{X,Y}(x, y) = P((X - x)(Y - y) > 0) - P((X - x)(Y - y) < 0) \quad (4)$$

For example if  $x = \text{med}(X)$ ,  $y = \text{med}(Y)$  than one obtain the  $q$ -measure of association of Kruskal (simplest measure of association in terminology by Kruskal). If  $x = E(X)$ ,  $y = E(Y)$  then one gets the sign correlation of Fehner [22]. In addition as it was proven by Lehmann if two random variables are positively dependent than its Kendall and Spearman correlations are positive. Therefore measures of association constitute a large family of measures of dependence between two random variables. In what follows we will use the notation  $\gamma_{X,Y}$  for any measure of association for two random variables  $X$  and  $Y$ .

### 3 Identification Problem in Market Network Analysis

We model the financial market as a family of random variables  $X_i(t)$ , where  $i = 1, 2, \dots, N$ ,  $t = 1, 2, \dots, n$ . In this setting  $N$  is the number of stocks and  $n$  is a number of observations. Random variable  $X_i(t)$  for a fixed  $i$ ,  $t$  describes the behavior of some numerical characteristic (price, return, volume and so on) of the stock  $i$  at the moment  $t$ . For a fixed  $i$  the sequence of random variables  $(X_i(1), X_i(2), \dots, X_i(n))$  describes the behavior of the stock  $i$  over the time. We assume that for a fixed  $i$  the random variables  $X_i(t)$  are independent and identically distributed as  $X_i$ . This assumption is valid for stocks returns and many other stocks characteristics. The random vector  $X = (X_1, X_2, \dots, X_N)$  gives a complete description of the market for the given numerical characteristic.

In this paper we consider only market network models based on the pair wise dependence of stocks. The nodes of the network are the stocks of the market and the weighted link between stocks  $i$  and  $j$ ,  $i \neq j$  is given by a measure of association  $\gamma_{i,j}$  for random variables  $X_i$  and  $X_j$ :  $\gamma_{i,j} = \gamma(X_i, X_j)$ . We call the obtained network *true* market network with measure of association  $\gamma$ . For a given structural characteristic  $S$  (MST, PMFG, MG, MC, MIS and others) *true* characteristic is obtained by filtration on the true market network. In general measure of association  $\gamma$  has to reflect a dependence between random variables associated with stocks. The choice of the measure of association is therefore connected with the joint distribution of the vector  $(X_1, X_2, \dots, X_N)$ . The most popular measure of association used in the literature is Pearson correlation. Pearson correlation is known to be the most appropriate measure of association in the case of multivariate normal distribution of the vector  $(X_1, X_2, \dots, X_N)$ . When the distribution of this vector is not known one needs a more universal measure of association not related with the form of distribution. One such measure of association is  $q$ -measure of Kruskal.

In practice however market networks are constructed from statistical data sets of observations. Let  $x_i(t)$  be an observation of the random variable  $X_i(t)$ ,  $i = 1, 2, \dots, N$ ,  $t = 1, 2, \dots, n$ . For a given structural characteristic  $S$  (MST, PMFG, MG, MC, MIS and others) the main problem is to identify true characteristic (associated with the true market network) from the observations. Traditional

way for this identification used in the literature can be described as follows: first one has to make estimations  $\hat{\gamma}_{i,j}$  of the measures of association  $\gamma_{i,j}$ , next one constructs the *sample* network as the weighted complete graph where the nodes are the stocks of the market and the weighted link between stocks is given by  $\hat{\gamma}_{i,j}$ . Finally, the structural characteristic  $S$  is identified on the sample market network by the same filtration process as on the true market network. Described identification process can be considered as statistical procedure for the identification of  $S$ . But this statistical procedure is not only one that can be considered for identification of  $S$ . Moreover it is not clear whether this procedure is the best possible or even if this procedure is good from statistical point of view. This question is crucial in our investigation.

## 4 Multiple Decision Theory

To answer the question above and define optimal statistical procedures for identification of structural characteristics one needs to formulate this problem in the framework of mathematical statistics theory. Identification of a given structural characteristic (MST, PMFG, MG, MC, MIS and others) is equivalent to the selection of one particular structural characteristic from the finite family of possible ones. Any statistical procedure of identification is therefore a multiple decision statistical procedure. Multiple decision theory is nowadays one of the active branch of mathematical statistics [12,20]. In the framework of this theory the problem of identification of FSC can be presented as follows. One has  $L$  hypothesis  $H_1, H_2, \dots, H_L$  corresponding to the family of possible subgraphs associated with FSC. Multiple decision statistical procedure  $\delta(x)$  is a map from the sample space of observations  $R^{N \times n} = \{x_i(t) : i = 1, 2, \dots, N; t = 1, 2, \dots, n\}$  to the decision space  $D = \{(d_1, d_2, \dots, d_L)\}$ , where  $d_j$  is the decision of acceptance of the hypothesis  $H_j$ ,  $j = 1, 2, \dots, L$ . Quality of the multiple decision statistical procedure  $\delta(x)$  according to Wald [25] is measured by its conditional risk. In our case conditional risk  $R(H_k, \delta)$  can be written as

$$R(H_k, \delta) = \sum_{j=1}^L w_{k,j} P_k(\delta(x) = d_j)$$

where  $w_{k,j}$  is the loss from the decision  $d_j$  when the true decision is  $d_k$ ,  $w_{k,k} = 0$ ,  $P_k(\delta(x) = d_j)$  is the probability to take the decision  $d_j$  when the true decision is  $d_k$ . Conditional risk can be used for the comparison of different multiple decision statistical procedures for structural characteristic identification [7] and it is appropriate to measure the statistical uncertainty of structural characteristics [6].

**Example 1. Market graph.** For a given value of threshold  $\gamma_0$  market graph [3] is obtained from the complete weighted graph (market network) by eliminating all edges with property  $\gamma_{i,j} \leq \gamma_0$ , where  $\gamma_{i,j}$  is the measure of association between stocks  $i$  and  $j$ . In this case the set of hypotheses is

$$\begin{aligned}
H_1 &: \gamma_{i,j} \leq \gamma_0, \forall (i,j), i < j, \\
H_2 &: \gamma_{12} > \gamma_0, \gamma_{i,j} \leq \gamma_0, \forall (i,j) \neq (1,2), i < j, \\
H_3 &: \gamma_{12} > \gamma_0, \gamma_{13} > \gamma_0, \gamma_{i,j} \leq \gamma_0, \forall (i,j) \neq (1,2), (i,j) \neq (1,3), \\
&\dots \\
H_L &: \gamma_{i,j} > \gamma_0, \forall (i,j), i < j,
\end{aligned} \tag{5}$$

where  $L = 2^M$  with  $M = N(N - 1)/2$ . These hypotheses describe all possible market graphs. To identify the true market graph one needs to construct a multiple decision statistical procedure  $\delta(x)$  which will select one hypothesis from the set  $H_1, H_2, \dots, H_L$ .

**Example 2. Minimum spanning tree (MST).** Minimum spanning tree [13] is the spanning tree of the complete weighted graph (market network) with the maximal total associations between included edges. In this case one has by Caylay formula  $L = N^{N-2}$  and each hypothesis  $H_s$  can be associated with multi-index  $s = (s_1, s_2, \dots, s_N, s_{N+1}, \dots, s_{2N})$ ,  $s_j \in \{0, 1\}$  (tree code).

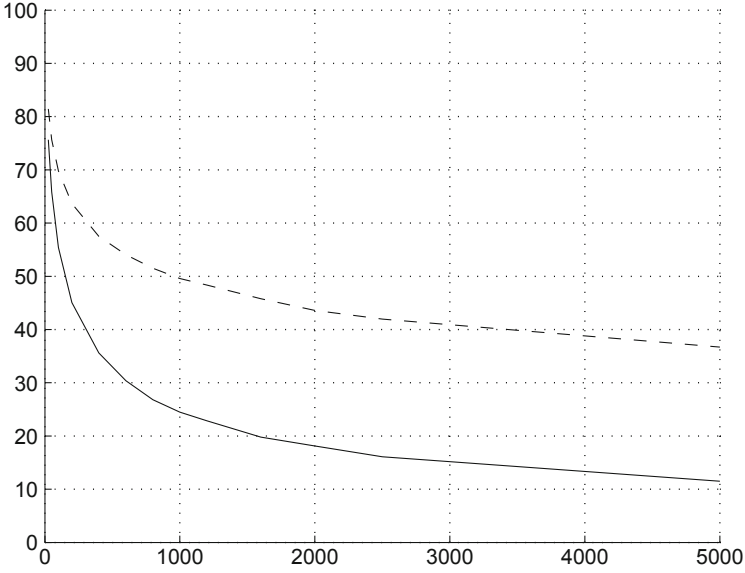
## 5 Conditional Risk

There are many ways to define the losses  $w_{k,j}$  and associated conditional risk. For example for a given structural characteristic  $S$  one can define a conditional risk by

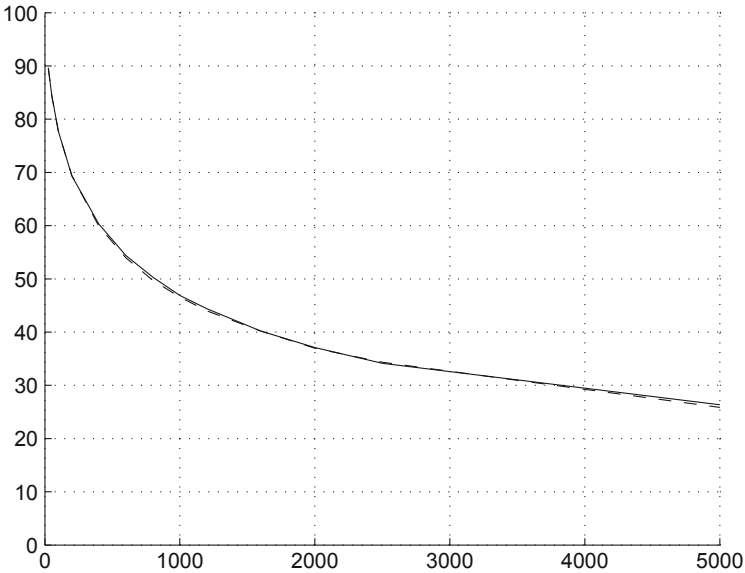
$$R(S, \delta) = \sum_{1 \leq i < j \leq N} [a_{ij} P_{i,j}^a(S, \delta) + b_{ij} P_{i,j}^b(S, \delta)], \tag{6}$$

where  $a_{i,j}$  is the loss from erroneous inclusion of the edge  $(i, j)$  in the structure  $S$ ,  $P_{i,j}^a(S, \delta)$  is the probability that decision procedure  $\delta$  takes this decision,  $b_{i,j}$  is the loss from erroneous non inclusion of the edge  $(i, j)$  in the structure  $S$ ,  $P_{i,j}^b(S, \delta)$  is the probability that decision procedure  $\delta$  takes this decision. Two terms in (6) can be considered as type I and type II statistical errors [12]. The value of conditional risk  $R(S, \delta)$  essentially depends on the choice of measure of association  $\gamma$ , distribution of random vector  $X = (X_1, X_2, \dots, X_N)$ , structural characteristic  $S$ , multiple decision statistical procedure  $\delta(x)$  for structural characteristic identification and number of observations  $n$ . To illustrate this dependence we present below some results of numerical experiments for MST on US stock market with  $N = 100$ ,  $a_{i,j} = b_{i,j} = 1/2$ . The experiments show some intriguing properties of associated conditional risk. The Fig. 1 shows the behavior of conditional risk for Pearson correlation, two type of distributions (multivariate Normal and Student distributions) and different number of observations. The Fig. 2 shows the behavior of conditional risk for Kruskal correlation, the same type of distributions (multivariate Normal and Student distributions) and different number of observations. In both cases the multiple decision statistical procedure is the Kruskal algorithm applied to the sample network (we use classical estimations for Pearson and Kruskal correlations).

The Fig. 1 shows that conditional risk for Pearson correlation has a big dependence on the type of distribution. Pearson correlation is a good measure of



**Fig. 1.** Conditional risk as a function of number of observations for Pearson correlation. Solid line corresponds to the normal distribution. Dashed line corresponds to the Student distribution.



**Fig. 2.** Conditional risk as a function of number of observations for Kruskal correlation. Solid line corresponds to the normal distribution. Dashed line corresponds to the Student distribution.

**Table 1.** Conditional risk for MST: Pearson correlation

n	Normal	Truncated normal	Platykurtic	Bimodal	Stable trend rare risk	Student
25	75.647	88.045	75.806	75.718	78.245	81.428
50	65.858	82.403	65.777	65.865	73.120	75.658
100	55.409	74.977	55.078	54.625	71.409	69.790
200	45.056	66.332	44.494	44.365	70.849	63.674
400	35.587	56.704	35.305	34.802	61.127	57.495
600	30.375	50.923	29.910	29.629	53.221	54.051
800	26.812	46.773	26.435	26.564	48.298	51.527
1000	24.468	43.402	24.172	23.957	44.119	49.588
1200	22.857	40.999	22.506	22.063	41.104	48.342
1600	19.799	37.111	19.814	19.495	36.419	45.798
2000	18.137	34.391	17.698	17.626	33.170	43.582
2500	16.099	31.651	15.830	15.829	30.139	41.946
5000	11.496	23.736	11.320	11.567	21.961	36.697

**Table 2.** Conditional risk for MST: Kruskal correlation

n	Normal	Truncated normal	Platykurtic	Bimodal	Stable trend, rare risk	Student
25	89.660	89.753	90.074	90.478	90.055	89.576
50	84.568	84.540	84.947	85.839	81.767	84.477
100	77.667	77.683	78.626	79.643	72.985	77.658
200	69.335	69.500	70.503	72.081	63.605	69.393
400	60.303	59.921	61.548	63.293	53.011	59.899
600	54.377	54.130	55.689	57.310	46.885	53.948
800	50.327	49.960	51.069	53.465	42.851	49.716
1000	46.903	46.583	48.068	50.129	39.664	46.599
1200	44.398	44.246	45.461	47.281	37.597	43.986
1600	40.209	40.187	41.552	43.365	33.948	40.190
2000	37.095	37.306	38.492	40.056	31.363	36.980
2500	34.138	34.482	35.452	37.153	28.858	34.345
5000	26.338	26.028	27.313	28.837	22.269	25.854

association for normal distribution and it is not good for Student distribution. The Fig. 2 shows that conditional risk for Kruskal correlation is stable with respect to the type of distribution. At the same time Kruskal correlation is more appropriate measure of association for Student distribution than Pearson



correlation. It suggests to use the Kruskal measure of association in the case of distributions with fat tails.

The values of conditional risk for different distributions and number of observations are presented in the Table 1 (Pearson correlation) and Table 2 (Kruskal correlation). All multivariate distributions in the tables have the same covariance matrix  $\Sigma$  (covariance matrix for the 100 stocks of US stock market) and are obtained by transformation  $X = \sigma^{1/2}Z$ ,  $Z = (Z_1, Z_2, \dots, Z_N)$  being the vector of normalized independent random variables with the same uni-variate distribution. This uni-variate distribution are normal, truncated normal, uniform distribution (platykurtic), distribution with two modes (bimodal), discrete distribution with 2 values (stable trend rare risk) and Student distribution with 3 degrees of freedom. Detailed description of these distributions is given in [1]. The Tables 1 and 2 confirm the stability of conditional risk for Kruskal correlation. A comparative analysis of conditional risk for Pearson and sign correlations for the market graph construction is given in [1] where some interesting observations are described. The problem of optimality of multiple decision statistical procedures for the market graph identification is discussed in [7]. It was proven in [7] that it is possible to construct a statistical procedures with lower conditional risk than the widely used in the literature statistical procedure based on the sample graph. The dependence of conditional risk on the filtered structural characteristic is investigated in [6].

## 6 Concluding Remark

The general approach to market network analysis for statistical data set gives an appropriate theoretical basis for investigation of different market network models. It allows to design a statistical procedures of a good quality for identification of structural characteristics of network.

## References

1. Bautin, G., Kalyagin, V.A., Koldanov, A.P.: Comparative analysis of two similarity measures for the market graph construction. In: Goldengorin, B.I., Kalyagin, V.A., Pardalos, P.M. (eds.) *Models, Algorithms, and Technologies for Network Analysis*. Springer Proceedings in Mathematics & Statistics, vol. 59, pp. 29–41. Springer, New York (2013)
2. Bautin, G.A., Kalyagin, V.A., Koldanov, A.P., Koldanov, P.A., Pardalos, P.M.: Simple measure of similarity for the market graph construction. *Comput. Manage. Sci.* **10**, 105–124 (2013)
3. Boginsky, V., Butenko, S., Pardalos, P.M.: On structural properties of the market graph. In: Nagurney, A. (ed.) *Innovations in Financial and Economic Networks*, pp. 29–45. Edward Elgar Publishing Inc., Northampton (2003)
4. Boginski, V., Butenko, S., Pardalos, P.M.: Statistical analysis of financial networks. *J. Comput. Stat. Data Anal.* **48**(2), 431–443 (2005)
5. Boginski, V., Butenko, S., Pardalos, P.M.: Mining market data: a network approach. *J. Comput. Oper. Res.* **33**(11), 3171–3184 (2006)

6. Kalyagin, V.A., Koldanov, A.P., Koldanov, P.A., Pardalos, P.M., Zamaraev, V.A.: Measures of uncertainty in market network analysis. *Physica A: Stat. Mech. Appl.* **413**, 59–70 (2014)
7. Koldanov, A.P., Koldanov, P.A., Kalyagin, V.A., Pardalos, P.M.: Statistical procedures for the market graph construction. *Comput. Stat. Data Anal.* **68**, 17–29 (2013)
8. Koldanov, A.P., Koldanov, P.A.: Optimal multiple decision statistical procedure for inverse covariance matrix. In: Demyanov, V.F., Pardalos, P.M., Batsyn, M. (eds.) *Constructive Nonsmooth Analysis and Related Topics. Springer Optimization and Its Applications*, vol. 87, pp. 205–216. Springer, New York (2014)
9. Koldanov, P.A.: Efficiency analysis of branch network. In: Goldengorin, B.I., Kalyagin, V.A., Pardalos, P.M. (eds.) *Models, Algorithms, and Technologies for Network Analysis. Springer Proceedings in Mathematics & Statistics*, vol. 59, pp. 71–83. Springer, New York (2013)
10. Kruskal, W.H.: Ordinal Measures of Association. *J. Am. Stat. Assoc.* **53**, 814–861 (1958)
11. Lehmann, T.L.: Some concepts of dependence. *Ann. Math. Stat.* **37**, 1137–1153 (1966)
12. Lehmann, E.L., Romano, J.P.: *Testing Statistical Hypotheses*. Springer, New York (2005)
13. Mantegna, R.N.: Hierarchical structure in financial market. *Eur. Phys. J. Ser. B* **11**, 193–197 (1999)
14. Newman, M.E.J.: *Networks: An Introduction*. Oxford University Press, New York (2010)
15. Newman, M.J.E., Barabasi, A.L., Watts, D.J.: *The Structure and Dynamics of Networks*. Princeton University Press, Princeton (2006)
16. Onnela, J.-P., Chakraborti, A., Kaski, K., Kertesz, K., Kanto, A.: Dynamics of market correlations: taxonomy and portfolio analysis. *Phys. Rev. E* **68**, 56–110 (2003)
17. Pardalos, P.M., Rebennack, S.: Computational challenges with cliques, quasi-cliques and clique partitions in graphs. In: Festa, P. (ed.) *SEA 2010. LNCS*, vol. 6049, pp. 13–22. Springer, Heidelberg (2010)
18. Rebennack S., Maximum Stable Set Problem: A Branch and Cut Solver, Ruprecht-Karls-Universitt Heidelberg, Fakultt fr Mathematik und Informatik (2006)
19. Shirokikh, J., Pastukhov, G., Boginski, V., Butenko, S.: Computational study of the US stock market evolution: a rank correlation-based network model. *Comput. Manage. Sci.* **10**(2–3), 81–103 (2013)
20. Rao, C.V., Swarupchand, U.: Multiple comparison procedures - a note and a bibliography. *J. Stat.* **16**, 66–109 (2009)
21. Shiryaev, A.N.: *Essentials of Stochastic Finance: Facts, Models, Theory. Advanced Series on Statistical Science and Applied Probability*. World Scientific Publishing Co., New Jersey (2003)
22. Stuart, A., Ord, J.K., Arnold, S.: *Kendalls Advanced theory of Statistics. Classical Inference and Relationships*, vol. 2A. Wiley, London (2004)
23. Tumminello, M., Aste, T., Matteo, T.D., Mantegna, R.N.: A tool for filtering information in complex systems. *Proc. Nat. Acad. Sci.* **102**(30), 10421–10426 (2005)
24. Tumminello, M., Lillo, F., Mantegna, R.N.: Correlation, hierarchies and networks in financial markets. *J. Econ. Behav. Organ.* **75**, 40–58 (2010)
25. Wald, A.: *Statistical Decision Function*. Wiley, New York (1950)