

1.9. ФОЛКСОНОМІЯ СОЦІАЛЬНО-ЕКОНОМІЧНИХ ОБ'ЄКТІВ В СКЛАДНИХ МЕРЕЖАХ ЗАСОБАМИ CORRANK

Останнім часом наукова спільнота анонсувала мережну парадигму як один з актуальних міждисциплінарних напрямків дослідження. Її суть зводиться до вивчення складних систем шляхом аналізу їх мережних властивостей, навіть в тих випадках, коли мережна структура не є самоочевидною [1,2].

Постановка проблеми. На початку XXI століття теоретично-методологічні основи дослідження та моделювання динамічних систем, були істотно розширені нелінійними методами, одержаними з теорії нелінійної динаміки і хаосу. Одним із таких методів є аналіз складних мережеподібних структур, сформованих на основі часових рядів. Цей метод дає можливість отримати додаткові характеристики часових рядів. Форма мережі притаманна багатьом системам, зокрема Інтернету, нейронним, телекомунікаційним, транспортним, соціальним, мережам господарських зв'язків різного рівня, тощо.

Елементи складної системи та їх взаємодії можна охарактеризувати за допомогою складної мережі. Мережею (*network*) називається сукупність вузлів (*node*), поєднаних зв'язками (*link*), яка будується за певними правилами. Складні мережі є об'єктом як теоретичних, так і емпіричних досліджень, в яких топологія мереж відіграє провідну роль [3].

Загальна концепція мережі є дуже інтуїтивною: мережа описує колекцію вузлів або вершин (наприклад, фінансові установи, фондові ринки) і зв'язки між ними, які можуть бути направленними (тобто, дуги) або ненаправленими (тобто, ребра).

Головна передумова аналізу мережі є те, що структура зв'язків між вузлами впливає на результати діяльності кожного вузла і системи в цілому. Властивості і поведінка конкретного вузла не можуть бути проаналізовані на основі своїх власних властивостей і поведінки поодиноці, так як вони можуть порушувати властивості вузлів, які мають зв'язки з ним, а також інших вузлів, які не мають прямих зв'язків, але пов'язані з його сусідами.

Перспективи розвитку складних мереж стимулювали вибувовий інтерес у дослідженнях соціальних, інформаційних, технологічних та біологічних систем, в результаті чого було досягнуто більш глибоке розуміння природи складних систем [4]. В рамках вивчення соціальних систем було проаналізовано мережні властивості багатьох економічних і фінансових систем. Активи компаній фондовому ринку, що належать до різних секторів економіки є корельованими. Якщо ми представимо кожен акцію у вигляді вузла, то «відстані» між двома акціями можна відобразити через коефіцієнти крос-кореляції. Оскільки вузли є взаємозалежними, то ринок формує мережу (граф). Мінімальне острівне дерево будується з матриці відстаней між активами та може бути використане для дослідження ієрархічної структури портфеля акцій [5,6].

Існують також методи трансформації немережних форм прояву (реалізації) складних систем у мережні. Так, не розглядаючи портфель акцій, часовий ряд ціни окремої акції також може бути відображеним у вигляді складної мережі, що дозволяє нам досліджувати динаміку акцій, використовуючи мережні засоби.

Складні фінансові та економічні мережі можна розділити на три категорії. У першій категорії мереж вузли представляються фінансовими чи економічними агентами (компанії, фінансові інститути, трейдери, тощо), зв'язаними ребрами, якщо вони мають певні взаємодії (наприклад, інвестиції, торгівля, кредитування, економічне співробітництво та ін.) [7]. У другій категорії мережі, вузли задані матрицею різних агентів, кожному з них відповідає часовий ряд, а динаміка поведінки зв'язків визначається кореляціями між парами агентів [8]. Третя категорія містить методи перетворення часових рядів у мережеподібні структури, серед яких найбільш застосовними є рекурентний метод [9] та граф видимості [10,11].

Багато в чому завдяки розвитку теорії складних систем, значно зросла наукова активність в області дискретної математики, що переживає період інтенсивного розвитку та проникнення в різні області наукового пізнання. Найбільш яскраво її впровадження проявилось в теорії графів, значні можливості якої обумовлені теоретико-множинними, комбінаторними та топологічними аспектами, що складають

основу поняття самого графу. Успіх застосування теорії графів можна пояснити також тим, що вона є зручною мовою для формулювання задач, які можна віднести до широкого кола наукових проблем, та зручним інструментом для їх вирішення.

Аналіз останніх досліджень та публікацій. В останнє десятиліття значний прогрес на шляху до розуміння структури та функціонування складних мереж було здійснено в галузі статистичної механіки та аналізу соціальних мереж. Зокрема, було розглянуто декілька мір центральності (centrality).

Найбільш часто використовуються такі міри центральності як ступеня (Degree), близькості (Closeness) та посередництва (Betweenness) [12]. Також часто використовують різні варіації вектор-центральності (Eigenvector centrality) [13,15,16].

Міра ступінь-центральності підраховує кількість сусідів кожного вузла. Це локальна міра, яка бере до уваги тільки безпосередню близькість від вузла. Вектор-центральність включає ідею про те, що центральне місце вузла безпосередньо залежить від центральності вузлів, які пов'язані з ним. Відомим алгоритмом вектор-центральності є алгоритм PageRank Google [17], який додає випадкові стрибки імовірностей "обірваних" (dangling) вузлів і таким чином дозволяє використовувати дану міру для всіх типів мереж. PageRank і Eigenvector centrality можна порівнювати з часткою часу, проведеного на відвідування кожного вузла при нескінченному випадковому блуканню по мережі.

Нещодавно розглянуто методи розрахунку центральності, створені для фінансової сфери [18,19].

Постановка завдання. На основі викладеного вище можна сформулювати завдання дослідження, котре полягає в розробці нової міри центральності, котра буде заснована на кореляційних відносинах елементів у побудованій мережі.

Виклад основного матеріалу дослідження. Для аналізу властивостей графу будемо виходити з матриці суміжності – квадратної матриці A розміру n , в якій значення елемента a_{ij} рівне числу ребер, які виходять з i -ї вершини графа в j -у вершину.

Для визначення кількісно кореляцій спочатку обчислюється зміна цін (прибутковості) акції $i = 1, \dots, N$ протягом часу Δt :

$$G_i(t) = \ln S_i(t + \Delta t) - \ln S_i(t),$$

де $S_i(t)$ позначає ціну акції i . Оскільки різні акції мають різні стандартні відхилення, визначатимемо нормалізовану прибутковість:

$$g_i(t) \equiv \frac{G_i(t) - \langle G_i \rangle}{\sigma_i},$$

де $\sigma_i \equiv \sqrt{\langle G_i^2 \rangle - \langle G_i \rangle^2}$ – стандартне відхилення G_i , а $\langle \dots \rangle$ позначає середнє значення за період часу, що досліджується. Тоді обчислення матриці взаємних кореляцій C проводиться згідно формули:

$$C_{ij} \equiv \langle g_i(t) g_j(t) \rangle.$$

За побудовою елементи C_{ij} обмежені областю $-1 \leq C_{ij} \leq 1$, де $C_{ij} = 1$ відповідає повним кореляціям, $C_{ij} = -1$ – повним антикореляціям, і $C_{ij} = 0$ свідчить про некорельованність пар акцій [20].

Наступним кроком обираємо пороговий критерій r переходу від кореляційної матриці до матриці суміжності. Для цього у якості відстані в полі кореляцій введемо величину, зв'язану із значенням коефіцієнта кореляції, середнє значення якої дорівнює $d_{ij}^{mean} = \sqrt{2(1 - C_{ij})}$. Тоді порогове значення буде знаходитись як величина, пропорційна цьому значенню: kd_{ij}^{mean} . В кожному конкретному випадку виникає необхідність підбору коефіцієнта k в залежності від природи складної системи. Наші дослідження показали, що оптимальним при розрахунках є значення $k = 1,1$. Таким чином, зв'язними є компанії, «відстань» між якими задовольняє критерію $r \leq kd_{ij}^{mean}$.

На рисунку 1 для матричної бази даних, до складу якої увійшли 80 підприємств індексної корзини S&P 500 побудовано різні види графів, у залежності від значення параметра k .

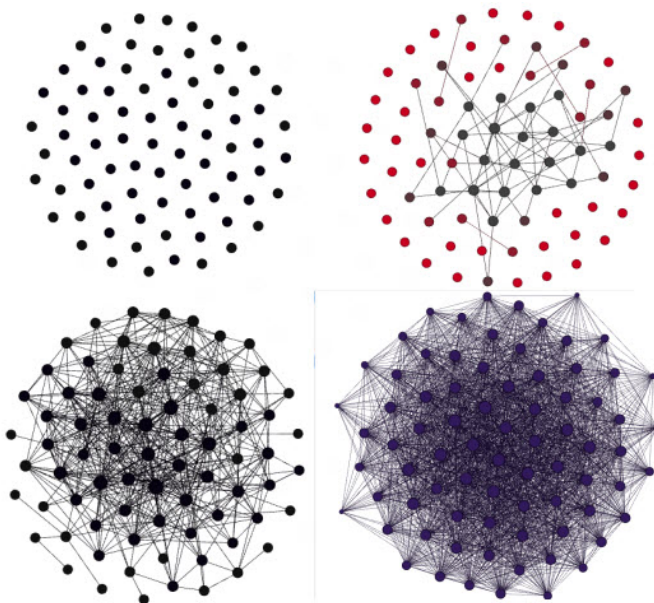


Рис.1 Візуалізований приклад графів, побудованих для матриці S&P 500 засобами Gerhi зі значеннями $k = (0,5; 0,75; 0,9; 1,1)$, відповідно.

Введемо локальний показник центральності графа – кореляційне ранжування (CorrRank).

CorrRank – це система рейтингування, ранжування та визначення важливості елементів складних систем, яка залежить від числа зовнішніх зв'язків і від їх ваги.

Представимо, що ми рухаємося по графу, вузлами якого є визначені компанії. Нехай почнемо з елемента p , випадкове блукання при цьому знаходиться в стані p . На кожному кроці ми проходимо наступний вузол, який має тісні кореляційні зв'язки з попереднім. Імовірність випадкового стрибка позначимо як d , тоді імовірність переходу по ребру буде $1 - d$. Таким чином, імовірність знаходження в даний момент на вершині p можна обчислити за формулою:

$$R_{j+1}(p) = d + (1 - d) \sum_{i=1}^k R_j(p_j) / L(p_j),$$

де $R(p)$ – значення міри CorRank елементу системи, $L(p_j)$ – кількість зв'язків між елементами (ребра графа), k – кількість зв'язків з елементом p в системі, d – коефіцієнт згасання (damping factor), що відображає частку ваги, яку може передати елемент, з якого виходить зв'язок, на елемент, котрий його приймає. Зазвичай його визначають рівним 0,85, що означає, що елемент може передати 85% ваги.

Якщо масштабувати CorRank таким чином, що

$$\sum_{i=1}^N R(p_i) = 1,$$

де N - число всіх елементів, для яких проводиться розрахунок CorRank, то $R(p)$ можна розглядати як розподіл ймовірності по всіх елементах (вершинах графа).

Для обчислення показника CorRank складається матриця M розміром $N \times N$, де кожному елементу M_{ij} матриці присвоюється значення $R_0(p)$ у тому випадку, якщо i -ий елемент має зв'язок з j -им, а всі інші елементи матриці заповнюються нулями. Таким чином, обчислення CorRank зводиться до відшукування власного вектора матриці M , що досягається множенням матриці M на вектор R_j на кожному кроці ітерації. Введення коефіцієнта згасання гарантує, що процес сходиться.

Наведемо приклад розрахунку CorRank для графа з трьома вершинами та зв'язками між ними, коефіцієнт згасання d візьмемо рівним 0,85. Як показали дослідження, саме значення 0,85 є оптимальним коефіцієнтом при аналізі складних мереж.

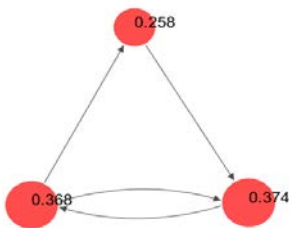


Рис.2. Розрахунок CorRank з коефіцієнтом згасання $d = 0,85$

Відмітимо, що при $d=1$ введений показник CorrRank буде відповідати мірі вектор-центральності. Для нашого прикладу при $d=1$, маємо для даного графа CorrRank рівним: 0.375, 0.375, 0.258 для трьох вершин відповідно.

Експериментальні результати. Нами була реалізована процедура розрахунку міри CorrRank в програмному середовищі Matlab (рис.3).

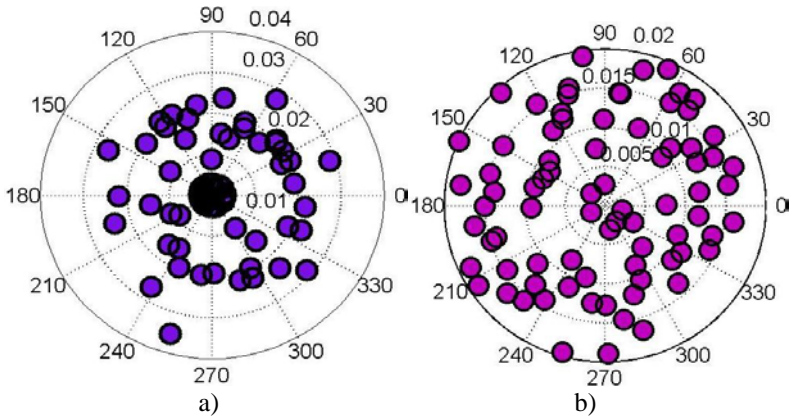


Рис.3 CorrRank для американського фондового ринку S&P500, розрахований для певного зрізу: а) спокійний стан; б) початок світової кризи 2008 року

З рисунку 3 можна бачити, що міра CorrRank реагує на кризові явища в системі, в тому числі характеризує взаємодію та зв'язність елементів (підприємств). Дана методика дає змогу прослідкувати, які з фірм є найбільш важливими, а також оцінити їх взаємозалежність. Помітно, що коли система знаходиться в спокійному стані (рис. 3а)), всі елементи знаходяться на певних позиціях та звичним чином пов'язані. Значення CorrRank в цьому випадку для більшості країн є близьким до нуля. Доречно відмітити, що найбільше значення CorrRank відповідає найбільш впливовим фірмам, а найменше мають суб'єкти найбільш віддалені та малозначущі на ринку. З рис. 3б) бачимо, що під час кризи всі елементи стають більш зв'язними та мають дуже близькі та дещо більші значення

CorrRank. Ті підприємства, котрі мають найбільше значення і є важливими суб'єктами, від яких найбільше залежить стабільне функціонування системи.

Введена міра центральності є так би мовити статичною характеристикою часових рядів. Але зрозуміло, що з плином часу в системі відбуваються зміни, котрі впливають на ступінь взаємодії елементів системи. Тому дослідимо ці зміни шляхом розрахунку введеної міри у рамках процедури ковзного вікна. В цьому випадку міра складності розраховується для підряду заданої довжини, після чого вікно зміщується («ковзає») у додатному напрямку і процедура повторюється до вичерпання значень часового ряду. Для зручності та спрощення аналізу дієвості введеної міри складності її зручно порівнювати з відповідною динамікою вихідного часового ряду.

Проведемо розрахунки віконної міри CorrRank для кожного підприємства з індексної корзини, що входить до створеної матричної бази даних фондових індексів США (S&P500) та Великобританії (FTSE).

Розрахунки проводилися для 80 підприємств індексу S&P500 та 47 підприємств з ринку FTSE за періоди: 01.04.1982 – 28.02.2014, 11.12.1989 – 12.02.2013 [21], відповідно, Було використано вікно в 250 днів та кроком 5 днів. Результати розрахунку впровадженої міри для вказаних даних зображено на рисунку 4.

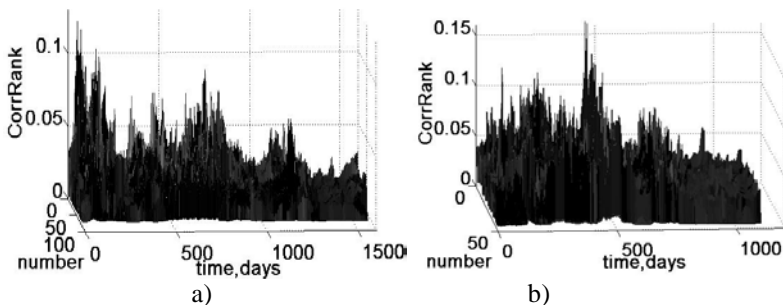


Рис.4 Динаміка міри CorrRank для індексних корзин фондових ринків: а) S&P500; б) FTSE

З рисунку 4, легко бачити певні спади та підвищення в динаміці міри CorrRank для кожного з вхідних елементів матриці на всіх поетапних кроках вікна..

Дослідимо, як реагує міра на кризові явища на ринку. При тому, що кожне підприємство з вхідної матриці для фондового ринку має свою певну динаміку, виникає необхідність усереднення всіх елементів.

Для зручності та спрощення аналізу дієвості міри складності її зручно порівнювати з відповідною динамікою вихідного часового ряду у відносних вимірниках, як це зроблено на рис. 5.

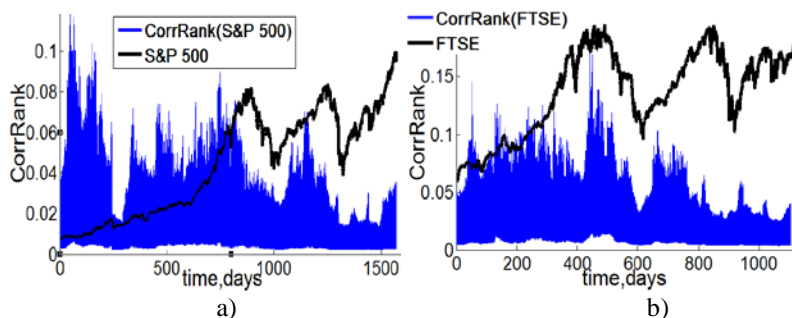


Рис.5 Порівняльна динаміка усередненої міри CorRank для елементів індексної корзини з фондовим індексом а) S&P500; б) FTSE[21]

Аналізуючи отриману динаміку, відмітимо, що введений показник можна інтерпретувати, як міру складності системи та проводити по ньому паралелі з кризовими явищами. Помітним зниженням міра CorRank реагує на кризові стани на ринку. В періоди релаксації міра зростає, що дає змогу використовувати її як адекватний індикатор-передвісник кризових та шоківих станів в складних фінансово-економічних системах.

Отже, у випадку усереднення всіх вхідних елементів створеної мережі у динаміці було показано, що показник CorRank знижується значно раніше настання кризових явищ на ринку. Проаналізуємо аналогічну міру для окремих компаній ринку (рис. 6).

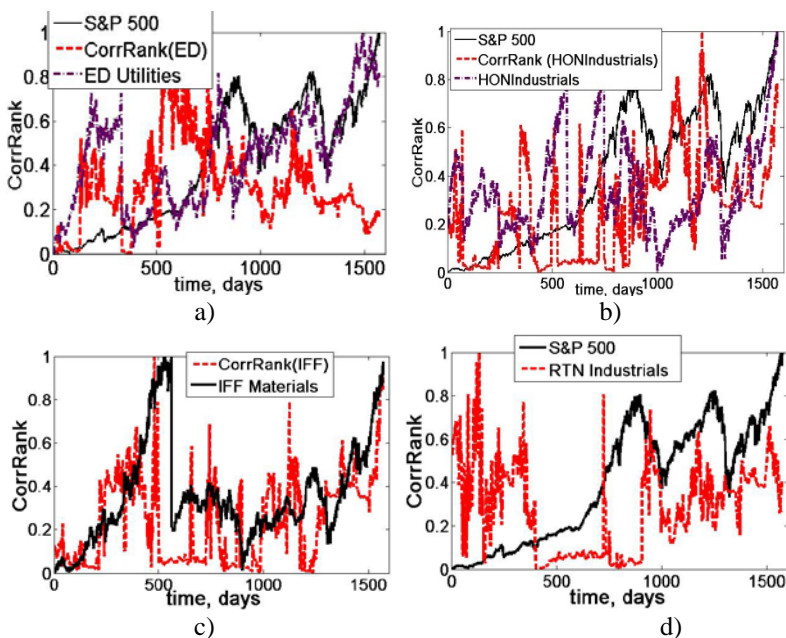


Рис. 6 Порівняльна динаміка міри кореляційного ранжування з вхідною динамікою [21]: а) американського фондового індексу з ED Utilities та CorrRank для нього; б) S&P 500 з динамікою підприємства HON Industrials та мірою CorrRank для другого; в) IFF Materials та розрахована для нього міра ранжування д) динаміки S&P 500 з CorrRank для підприємства RTN Industrial

З рисунка 6 видно, що динаміка для окремих підприємств з індексної корзини в різних випадках корелює між собою та з індексом S&P 500. Отже і в цьому випадку міра CorrRank виконує функції моніторингу та попередження кризових явищ в системі, знижуючись дещо раніше їх настання, що дає змогу використовувати даний показник як індикатор-передвісник несприятливих явищ у системах подібного типу.

Висновки з проведеного дослідження. Проведене дослідження дозволило розробити та апробувати нову методику кореляційного ранжування елементів в складних мережних системах, як в певний зріс часу, так і в динаміці.

Експериментальні дослідження показали універсальність та дієвість введеної міри складності CorrRank, як для ринка в загальному, так і для будь-якого з вхідних підприємств.

Показано, що дана міра попереднім зниженням реагує на наближення кризи в системі, що дає змогу для її використання як адекватного індикатора-передвісника несприятливих явищ в складних системах.

В подальших дослідженнях планується розгляд та апробація інших методик центральності графа.

Список використаної літератури

- 1.Zhang J. Complex Networks from Pseudoperiodic Time Series: Topology versus Dynamics / J. Zhang, M. Small // Boston. – 2006. – P. 35-42.
- 2.Voccatti S. Complex Networks: Structure and Dynamics / S. Voccatti, V. Latora, Y. Moreno, M. Chavez, D.-U. Hwang // Physics Reports. – 2006. – P. 153.
- 3.Хаханян В.Х. Елементи теорії складності алгоритмів // Навч. посіб. для студентів спеціальності АКБ. Видання друге, доповнене і перероблене. – М.: МІІТ.- 2010. – С. 95-96.
- 4.Newman M. E. J. The structure and function of complex networks // SIAM Reviews. – 2003. – V.45(2). – P.167-256.
- 5.Bonanno G. Taxonomy of stock market indices / G. Bonanno, N. Vandewalle, R.N. Mantegna // Phys. Rev. E 62, R7615.-2000.
- 6.Onnela J.-P. Dynamic asset trees and portfolio analysis / J.-P. Onnela, A. Chakraborti, K. Kaski, J. Kert'esz // Eur. Phys. J. B 30.- P. 285–288.
- 7.Kogut B. Emergent properties of a new financial market: American venture capital syndication / B. Kogut, P. Urso, G. Walker // [Електронний ресурс] – Режим доступу: doi:10.1287/mnsc.1060.0620.
- 8.Tumminello M. Correlation based networks of equity returns sampled at different time horizons / M. Tumminello, T. Di Matteo, T. Aste, R. N. Mantegna // Eur.Phys. J. B 55.- 2007.-P.209-217.
- 9.Donner R.V. Recurrence networks – A novel paradigm for nonlinear time series analysis / R.V. Donner, Y. Zou, J.-F. Donges, N. Marwan, J. Kurths // Eur. Phys. J. Spec. Top. 164. – 2008. – P. 28.
- 10.Lacasa L. From time series to complex networks: The visibility graph / L. Lacasa, B. Luque, F. Ballesteros, J. Luque // Proc. Natl. Acad. Sci. U.S.A. 105.-2008.- 4972–4975.
- 11.Xu X.-K. Superfamily phenomena and motifs of networks induced from time series / X.-K. Xu, J. Zhang, M. Small // Proc. Natl.

- Acad. Sci. U.S.A. 105.-2008.- P. 19601–19605.
doi:10.1073/pnas.0806082105
12. Borgatti S. P. / Centrality and network flow // *Social Networks*, 27: P.55–71.- [Електронний ресурс] – Режим доступу: dx.doi.org/10.1016/j.socnet.2004.11.008.
13. Freeman, L. Centrality in social networks: I. Conceptual clarification // *Social Networks* 1.- 1979.- P.215-239 [Електронний ресурс] – Режим доступу: <http://www.cin.ufpe.br/~rbcp/taia/Freeman1979-centrality.pdf>
14. Katz L. A new index derived from sociometric data analysis. *Psychometrika* // [Електронний ресурс] – Режим доступу: <http://ideas.repec.org/a/spr/psycho/v18y1953i1p39-43.html>.- 1953.
15. Bonacich P. Factoring and weighting approaches to status scores and clique identification // *Journal of Mathematical Sociology* .- 1972 .- 2(1): 113–120.
16. Bonacich P. Power and centrality: A family of measures // *American Journal of Sociology*.- 1978.- 92(5): P. 1170–1182.
17. Page L. The PageRank Citation Ranking: Bringing Order to the Web // L. Page, S. Brin, R. Motwani, T. Winograd // 1999-66 .- [Електронний ресурс] – Режим доступу: <http://ilpubs.stanford.edu:8090/422/>.
18. Battiston S. DebtRank: Too central to fail? Financial networks, the FED and systemic risk / S. Battiston, M. Puliga, R. Kaushik, P. Tasca, G. Caldarelli // [Електронний ресурс] – Режим доступу: doi:10.1038/srep00541 .- 02 August 2012.
19. Soramäki K. SinkRank: An Algorithm for Identifying Systemically Important Banks in Payment Systems/ K. Soramäki and S. Cook // *Economics: The Open-Access, Open-Assessment E-Journal*.- Vol. 7.- 2013-28.
20. Plerou V. Random matrix approach to cross correlations in financial data / V. Plerou, P. Gopikrishnan, B. Rosenow, L.A.N. Amaral, T. Guhr, H.E. Stanley // *Phys.Rev.E* – v.65, N 12.- 2002.- P.356-373.
21. Джерело статистики індексів світового фондового ринку [Електронний ресурс] – режим доступу: <http://finance.yahoo.com>