فصلنامــه اقتصاد مالی
دورهٔ ۱۷ / شمارهٔ ۱ (پیاپی ٦٢)/ بهار ۱٤۰۲
صفحه ۲۳۹ تا ۲۵۲
https://doi.org/10.30495/fed.2023.1934542.2455

داده‌کاوی بازار سهام ایران با مدل‌سازی فیلترینگ شبکه‌های پیچیده: رویکرد MST

تاریخ دریافت: ۱٤۰۱/۱۰/۰۵ تاریخ پذیرش: ۱٤۰۱/۱۲/۰۷

هادی اسماعیل‌پورمقدم[1]✉

چکیده

یکی از مهم‌ترین مسائل در مالی نوین، یافتن روش‌های کارآمد برای خلاصه کردن و تجسم داده‌های بازار سهام است. مدل‌سازی فیلترینگ شبکه‌های پیچیده در بازار سهام، این امکان را از طریق کاهش اندازه بازار، با دستیابی به اطلاعات قابل اطمینان و با اختلال کمتر فراهم می‌آورد. به‌علاوه، از آن‌جایی که تغییرات قیمت سهام مستقل از یکدیگر نیستند، مطالعه همبستگی تغییرات قیمت سهام با شبکه‌های پیچیده، درک بیشتری از عملکرد بازار برای سرمایه‌گذاران فراهم می‌نماید. در این مقاله، با استفاده از داده‌های بازار سهام در بورس اوراق بهادار تهران، شبکه بازار سهام ایران با روش آستانه ایجاد می‌شود و سپس فیلترینگ شبکه بر اساس می‌نیمم درخت فراگیر (MST) صورت می‌گیرد. نتایج نشان می‌دهد مدل‌سازی فیلترینگ شبکه بازار سهام ایران بر اساس می‌نیمم درخت فراگیر، می‌تواند زیرمجموعه‌ای از بازار سهام را تشکیل دهد که عملکرد کل بازار را با کاهش قابل توجهی در اندازه دنبال نماید و از درجه تنوع‌سازی مشابهی با کل بازار برخوردار باشد. نتایج تحقیق دلالت بر این دارد روش حاصل از فیلترینگ شبکه مبتنی بر MST، می‌تواند مجموعه سهام تقلیل یافته‌ای را نسبت به کل بازار ارائه دهد که رفتار کل بازار را منعکس می‌کنند. از این رو، به جای تحلیل داده‌های کل بازار سهام، می‌توان رفتار مجموعه سهام حاصل را بررسی نمود. این تحلیل‌ها امکان بینش عمیق‌تر ساختار داخلی بازار سهام را ضمن کاهش ابعاد فراهم می‌نماید.

واژه‌های کلیدی: فیلترینگ بازار سهام، تحلیل شبکه‌های پیچیده، می‌نیمم درخت فراگیر، متنوع‌سازی.
طبقه بندی JEL: D53, G11, G20

۱- گروه اقتصاد، دانشگاه فردوسی ،مشهد، ایران(نویسنده مسئول)، esmaeilpour@um.ac.ir

نوع مقاله: علمی
Ecj@iauctb.ac.ir

This work is licensed under a Creative Commons Attribution 4.0 International License

۱- مقدمه

میزان داده‌های روزانه تولید شده توسط بازار سهام، بسیار زیاد است؛ این اطلاعات معمولاً توسط هزاران نمودار نمایش داده می‌شود که منعکس کننده قیمت هر سهام در طی یک دوره زمانی مشخص می‌باشد. تحلیل این نمودارها با افزایش تعداد سهام، پیچیده‌تر می‌شود (بوگینسکی و همکاران[۱]، ۲۰۰۶). هم‌چنین وجود تعداد زیادی از اجزای ناهمگون، منجر به رفتار پیچیده بازار سهام می‌گردد (دیمتریوس و وسیلیوس[۲]، ۲۰۱۵). از این رو، روش‌های داده‌کاوی با خلاصه کردن و کاهش بعد داده‌های بازار سهام، امکان کسب اطلاعات مفید از بازار سهام را فراهم می‌نماید. اهمیت بالای تکنیک‌های داده‌کاوی برای کشف دانش از پایگاه‌های داده غیر قابل انکار است (خطیب سمنانی و همکاران، ۱۳۹۶ و ولایتی و همکاران، ۱۳۹۶). از سوی دیگر، تغییرات قیمت سهام مستقل نیستند و با بخش‌های کسب و کار و صنایعی که سهام متعلق به آن‌ها هستند، همبستگی قوی و پیوستگی زیادی دارند (تیسی و همکاران[۳]، ۲۰۱۰). یکی از روش‌های داده‌کاوی، رویکرد شبکه‌های پیچیده و مدل‌سازی فیلترینگ است که برای خلاصه‌سازی داده‌های بازار سهام و مطالعه همبستگی قیمت‌های سهام پیشنهاد شده است (نیر و همکاران[۴]، ۲۰۰۸ و جلو و همکاران[۵]، ۲۰۱۳). شایان ذکر است که تجزیه و تحلیل همبستگی در امور مالی غیر معمول نیست. در واقع، تجزیه و تحلیل همبستگی در تخصیص دارایی و مدیریت ریسک به طور گسترده مورد مطالعه قرار گرفته است و منجر به نتیجه‌گیری مدل‌های بسیاری از مدیریت سبد و تخصیص دارایی شده است که با موفقیت در برنامه‌های دنیای واقعی مورد استفاده قرار می‌گیرند (تیسی و همکاران، ۲۰۱۰).

علاقه به استفاده از روش شبکه‌های پیچیده در تحقیقات مالی به‌ویژه بازار سهام به تازگی توسعه یافته است. تحلیل شبکه می‌تواند باعث افزایش تفهیم نوسانات سیستم‌های مالی شود (پکروآ و اسپلتا[۶]، ۲۰۱۵). تحلیل شبکه به عنوان ابزار متدولوژی قدرتمند برای مدل‌سازی تعاملات بین واحدهای اقتصادی، شرکت‌ها و نهادهای مالی شناخته می‌شود.

رویکرد تحلیل شبکه‌های پیچیده در بازار سهام می‌تواند تصویر روشنی از ساختار داخلی بازار سهام ارائه دهد (دیمتریوس و وسیلیوس، ۲۰۱۵). در این روش، برخلاف مدل‌های کلاسیک مبتنی بر هزینه-فایده، تغییرات قیمت سهام تحت تأثیر رفتارهای گروهی قرار می‌گیرد. مطالعه ساختار شبکه بازار سهام، رفتار بازار سهام و تعامل عوامل آن را با یکدیگر توضیح می‌دهد. بدین ترتیب این رویکرد، فرض در نظر گرفتن متغیر مستقل در روش‌های تحلیل خطی رایج مبتنی بر شناسایی نحوه تاثیر متغیر وابسته توسط چندین متغیر مستقل را به چالش می‌کشد (خیرخواه و همکاران[۷]، ۲۰۱۶). برای کاهش ازدحام داده‌های بازار سهام و دست‌یابی به اطلاعات قابل اتکا، نیاز به روشی

[1] . Boginski et al.
[2] . Dimitrios & Vasileios
[3] . Tse et al.
[4] . Nier et al.
[5] . Jallo et al.
[6] . Pecora & Spelta
[7] . Kheyrkhah et al.

کارآمد برای شبکه بازار سهام است که یال‌های[1] اضافی از شبکه بازار سهام حذف شود. به طور کلی، فرایندهای فیلترسازی که در ادبیات شبکه استفاده می‌شود، اغلب با فیلتر اتصال‌ها یا یال‌ها به حذف اختلال از ماتریس همبستگی شبکه بازار اقدام می‌کنند تا ماتریس همبستگی را قابل اتکاتر برای بهینه‌سازی سبد نمایند. روش می‌نیمم درخت فراگیر (MST[2]) یکی از معمول‌ترین و شناخته شده‌ترین این روش‌ها است که درختی پوشا از گرافی متصل بدون جهت را شکل می‌دهد؛ به طوری که جمع اوزان یال‌های انتخاب شده حداقل باشد. از این رو، این مطالعه در تلاش است با ایجاد شبکه‌ای کامل از ارتباطات مبتنی بر همبستگی برای سهام‌های بازار بورس ایران، اطلاعات مربوط به ساختار داخلی بازار سهام را منعکس نماید. سپس با مدل‌سازی فیلترینگ شبکه بر اساس MST، یال‌های اضافی را حذف نموده و زیرمجموعه‌ای کاهش یافته از شبکه بازار سهام را ارائه دهد؛ به طوری که عملکرد کل بازار را با کاهش قابل توجهی در اندازه دنبال نماید و از درجه تنوع‌سازی مشابهی با کل بازار برخوردار باشد. مزیت این رویکرد، در کاهش هزینه و ریسک بررسی داده‌های بزرگ بازار با رصد گروه نسبتاً اندکی از سهام به جای کل بازار است. دیگر مزیت مدل‌سازی فیلترینگ شبکه بر اساس MST این است که با حذف یال‌های اضافی، امکان درک شهودی آسان‌تر کل شبکه فراهم می‌آید که در این تحقیق نیز بدان پرداخته می‌شود. این نوع نمایش شهودی، راهی بصری برای تشخیص عملکرد سهام و رفتار مشابه آن‌ها ارائه می‌دهد. ادامه این مقاله به شرح زیر می‌باشد: بخش دو، ادبیات موضوع و مرور مطالعات قبلی در زمینه شبکه‌های پیچیده است. سپس بخش سه، روش تحلیلی را توصیف می‌کند و در نهایت، به ارائه نتایج و نتیجه‌گیری پرداخته می‌شود.

۲. ادبیات نظری و مرور پیشینه پژوهش

تحلیل همبستگی اهمیت قابل توجهی در مدیریت ریسک و تخصیص دارایی دارد (التون و گروبر[3]، ۱۹۹۵). مطالعه ماتریس‌های همبستگی دارای تاریخچه‌ای طولانی در امور مالی است و سنگ‌بنای اصلی نظریه مارکوویتز در مورد اوراق بهادار است (لالوکس و همکاران[4]، ۲۰۰۰). شبکه بازار سهام مطالعه همبستگی تغییرات قیمتی سهام را فراهم می‌آورد. استفاده از ابزار شبکه به منظور تجزیه و تحلیل وابستگی‌های اقتصادی یا به هم پیوستگی‌های اقتصاد بعد از بحران مالی جهانی ۲۰۰۹-۲۰۰۸ گسترش پیدا کرد که ماهیت شبکه بودن سیستم‌های بانکی را به عنوان مجموعه‌ای از شرکت‌ها که با ارتباط نزدیک با یکدیگر عمل می‌کنند، آشکار کرد. رویکرد شبکه مربوط به ساختار و تشکیل سیستم گره‌ها[5] می‌باشد. یک فرض بنیادی این است که تماماً ویژگی‌ها و رفتار یک گره می‌تواند تنها با توجه به ارتباط آن با بقیه سیستم مورد ارزیابی قرار گیرد. در شبکه‌های مالی، گره‌ها معمولاً مؤسسات مالی یا نهادهای مشابه هستند. هنگامی که عوامل با توجه به هزینه و منافع ضمنی، متصل و مرتبط می‌شوند، این فرآیند اتصال لزوماً بستگی به موقعیت نسبی آن‌ها در شبکه دارد (آلن و بابوس[6]، ۲۰۰۹).

[1]. Edge
[2]. Minimum Spaning Tree
[3]. Elton & Gruber
[4]. Laloux et al.
[5]. Node
[6]. Allen & Babus

پژوهش‌های متعددی پیرامون تحلیل شبکه در بازار سهام انجام گرفته است. کارایانی[۱] (۲۰۱۲)، یانگ و همکاران[۲] (۲۰۱۴)، کلتی[۳] (۲۰۱۶)، مجاپا و گسل[۴] (۲۰۱۶)، بریدا و همکاران[۵] (۲۰۱۶)، ژونگ و همکاران[۶] (۲۰۱۶)، ژائو و همکاران[۷] (۲۰۱۶)، ژنگ و همکاران[۸] (۲۰۱۷) و لانگ و همکاران[۹] (۲۰۱۷) پس از تحلیل ساختار توپولوژیکی شبکه بازارهای مالی، تحلیل شبکه را به عنوان راهنمایی مفید برای سرمایه‌گذاران معرفی نمودند. ابرهارد و همکاران[۱۰] (۲۰۱۷) در تحقیقی به بررسی ویژگی‌های شبکه‌ای بازار سهام شیلی پرداختند. نتایج این تحقیق نشان داد ساختار شبکه بازار سهام شیلی می‌تواند در بازده و حجم معاملات سهام در بازار مؤثر باشد. در مطالعه شرما و همکاران[۱۱] (۲۰۱۷) همبستگی سهام با استفاده از روش آستانه انجام گرفت و شبکه بازار سهام هند بدین مبنا تشکیل شد. نتایج حاکی از این بود که تجزیه و تحلیل شبکه‌ای بازار سهام هند، می‌تواند فهم بهتری از وابستگی‌های سهام در بازار سهام هند ارائه دهد. جورج و چنگات[۱۲] (۲۰۱۷) نیز در مطالعه‌ای از رویکرد تحلیل شبکه برای داده‌کاوی بازار سهام و تحلیل سبد استفاده نمودند. در این تحقیق با استفاده از معیارهای شبکه سهام‌های مؤثر و با نفوذ بالا شناسایی گردید. یافته‌های تحقیق نشان داد تحلیل شبکه داده‌های سهام می‌تواند نقش مهمی در مطالعه بازار سهام داشته باشد. اسماعیل‌پورمقدم و همکاران[۱۳] (۲۰۱۹) به بررسی نقش مرکزیت در بازار سهام ایران با استفاده از تحلیل شبکه‌های پیچیده پرداختند. نتایج مطالعه حاکی از این بود که مرکزیت در شبکه بازار سهام ایران، می‌تواند دلالت‌های مهم اقتصادی و مالی داشته باشد. بررسی‌های دیگر در زمینه مطالعه شبکه بازار سهام حاکی از این بود که رشد بخشی می‌تواند در همبستگی سهام‌ها در شبکه منعکس شود و بر اساس همبستگی سهام‌ها در شبکه، می‌توان شاخصی را ارائه نمود که ضمن کاهش بعد در قیاس با شاخص کل بورس، رفتار بازار را به خوبی منعکس نماید (اسماعیل‌پورمقدم و همکاران، ۱۳۹۸).

مدل‌سازی فیلترینگ شبکه بازار سهام نقش ویژه‌ای را در دست‌یابی به اطلاعات ضروری از شبکه بازار سهام ایفا می‌نماید. در این میان، فیلترسازی بر اساس MST به عنوان موضوعی مهم و پرکاربرد در شبکه مطرح شده است (ام و همکاران[۱۴]، ۲۰۰۹). پژوهش‌های مختلف نظیر کرولهو و گوپتا[۱۵] (۲۰۱۸)، جو و همکاران[۱۶] (۲۰۱۸)،

[1] . Caraiani
[2] . Yang et al.
[3] . Coletti
[4] . Majapa & Gossel
[5] . Brida et al.
[6] . Zhong et al.
[7] . Zhao et al.
[8] . Zhang et al.
[9] . Long et al.
[10] . Eberhard et al.
[11] . Sharma et al.
[12] . George & Changat
[13] . Esmaeilpour Moghadam et al.
[14] . Eom et al.
[15] . De Carvalho & Gupta
[16] . Jo et al.

تنگ و همکاران[۱] (۲۰۱۸)، شی و همکاران[۲] (۲۰۱۹)، سانگ و همکاران[۳] (۲۰۱۸) پیرامون MST، اهمیت این رویکرد را درشبکه بازار سهام تأیید می‌نماید. تمینلو و همکاران[۴] (۲۰۱۰) از MST به عنوان ابزاری برای فیلتر ارتباطات بین سهام استفاده کردند و خوشه‌بندی سلسله مراتبی را برای شبکه بازار سهام اعمال نمودند. فیدر[۵] (۲۰۱۴) با اعمال فیلترینگ MST بر شبکه داده‌های بازار سهام نزدک نشان داد این رویکرد منجر به نتایج مفیدی برای گسترش پویایی‌های بازار سهام می‌گردد. در مطالعه بیرچ و همکاران[۶] (۲۰۱۵) نیز تأیید گردید که مدل‌سازی فیلترینگ شبکه بازار سهام موجب ارائه بینشی جدید پیرامون پویایی‌های اقتصاد می‌گردد.

۳- روش‌شناسی پژوهش

تئوری شبکه‌های پیچیده، برگرفته از تئوری گراف می‌باشد که به عنوان یک چارچوب نظری برای درک ویژگی‌های ساختاری شبکه‌ها توسعه یافته است (کیتو و ادا[۷]، ۲۰۱۴). شبکه، یک سه تایی $G = (V, E, f)$ است که در آن V مجموعه‌ای متناهی از گره‌ها و $E \subseteq V \otimes V = \{e_1, e_2, \dots, e_m\}$ مجموعه‌ای از پیوندها یا یال‌ها و f نگاشتی است که برخی از عناصر E را به یک جفت از عناصر V مربوط می‌کند؛ به طوری که اگر $v_i \in V$ و $v_j \in V$ باشد، آن‌گاه خواهیم داشت: $f: e_p \longrightarrow [v_i, v_j]$ و $f: e_q \longrightarrow [v_j, v_i]$. اگر $p_i(t)$ قیمت پایانی سهم i در روز t باشد، بازده سهام در روز tام به صورت ذیل است:

$$r_i(t) = lnp_i(t) - \ln p_i(t-1) \qquad (۱)$$

همبستگی بین دو سری بازده سهام، یال اتصالی دو گره یا دو سهم مشخص می‌کند که به عنوان یک جمله از ماتریس همبستگی C در نظر گرفته می‌شود:

$$c_{ij} = \frac{\langle r_i r_j \rangle - \langle r_i \rangle \langle r_j \rangle}{\sqrt{\left(\langle r_i^2 \rangle - \langle r_i \rangle^2\right)\left(\langle r_j^2 \rangle - \langle r_j \rangle^2\right)}} \qquad (۲)$$

که r بازده و کروشه، میانگین زمانی در طول دوره است. علاوه بر این، مقدار آستانه معین θ، $0 \leq \theta \leq 1$ بر اساس متدولوژی ارائه شده توسط زو و همکاران[۸] (۲۰۱۸) تعیین می‌گردد و یک یال بدون جهت بین گره‌های i و j رسم می‌شود اگر مقدار قدرمطلق c_{ij} بزرگتر یا مساوی θ باشد. زو و همکاران (۲۰۱۸) بر این اساس که تغییر در شبکه ساخته شده باید با تغییر در بازار حقیقی سازگار باشد، یک تابع سازگار بین آن‌ها معرفی می‌کنند و آستانه بهینه

[1] . Tong et al.
[2] . Shi et al.
[3] . Song et al.
[4] . Tumminello et al.
[5] . Fiedor
[6] . Birch et al.
[7] . Kito & Ueda
[8] . Xu et al.

با توجه به ماکزیمم کردن سازگاری انتخاب می‌شود. بر این اساس، برای هر مقدار $\theta_i = [0,1]$، ماتریس همبستگی C_i و شبکه N_i ساخته می‌شود و تابع ذیل محاسبه می‌گردد:

$$G_{\theta_i} = \frac{\langle D_C D_N \rangle - \langle D_C \rangle \langle D_N \rangle}{\sigma_{D_C} \sigma_{D_N}} \quad (۳)$$

این تابع، سازگاری بین تغییرات در ماتریس همبستگی و شبکه را اندازه‌گیری می‌نماید که در آن، D_C و D_N به ترتیب تغییرات در ماتریس و شبکه و σ_{D_C} و σ_{D_N} مربوط به انحراف معیار آن‌ها است که با روش عددی مقدار آستانه بهینه بر اساس ذیل انتخاب می‌شود:

$$\hat{\theta} = arg\, \max_{\theta_i} \{G_{\theta_i}\} \quad (۴)$$

هم‌چنین مطابق با ژنگ و همکاران[1] (۲۰۱۰)، از قدرمطلق همبستگی به عنوان وزن اختصاص داده شده به یال در شبکه استفاده می‌گردد. در این صورت، $G = (V, E, W)$ نشان دهنده شبکه سهام است که V مجموعه‌ای از رأس‌ها را نشان می‌دهد، E نشان دهنده یال ها و W وزن یال است. W به شرح زیر تعریف می‌شود:

$$W = \begin{cases} w_{ij} = |c_{ij}|, & i \neq j \ and \ |c_{ij}| \geq 0 \\ w_{ij} = 0, & else. \end{cases} \quad (۵)$$

اگر $w_{ij} \neq 0$، آن‌گاه یک یال بین گره‌های i و jوجود خواهد داشت.

گام دوم پس از ایجاد شبکه بازار سهام، فیلترسازی شبکه با استفاده از می‌نیمم درخت فراگیر (MST) است. MST ابزاری شناخته شده در ادبیات شبکه برای فیلترسازی یال‌های اضافی از شبکه است. بدین منظور چنانچه درخت $T(V, E')$ به عنوان زیرمجموعه‌ای از $G(V, E)$ و $E' \subset E$ باشد، برای هر یال $e_{ij} = (v_i, v_j)$ در این درخت، وزن متناظر w_{ij} تعریف می‌شود؛ به طوری که $e_{ij} \in E$ و $v_i, v_j \in V$. درخت یکتای T برای گراف G، MST است اگر و فقط اگر مقدار تابع $\sum_{e \in E} w(e)$ می‌نیمم باشد.

الگوریتم‌های حل MST توسط بزلامسی و هندی[2] (۲۰۰۱) بررسی شده و بر اساس اندازه شبکه و عملکرد زمانی به دو طبقه‌بندی کلاسیک و مدرن تقسیم شده‌اند. الگوریتم‌های مدرن که اخیراً توسعه یافته‌اند، از روش‌های جستجوی تصادفی برای حل MST استفاده می‌کنند که باعث می‌شود مدل با سرعت بیشتری نسبت به الگوریتم‌های کلاسیک اجرا شود. از این رو، الگوریتم‌های مدرن پیچیده‌تر هستند؛ اما برای مسائل با مقیاس بزرگ کارآمدتر از الگوریتم‌های کلاسیک هستند. در این تحقیق، از الگوریتم کروسکال[3] به عنوان الگوریتمی مدرن برای حل MST استفاده می‌شود. الگوریتم کروسکال، یک الگوریتم حریصانه است که در آن یال‌ها بر اساس وزن به طور صعودی مرتب می‌شوند. در ابتدا یال e_1 با حداقل وزن انتخاب شده و سپس به ترتیب یال‌های دیگر، چنانچه نقطه پایانی

[1]. Zhang et al.
[2]. Bazlamacci & Hindi
[3]. Kruskal

یال در همان درخت باشد، انتخاب می‌گردند؛ در غیر این صورت، آن یال نادیده گرفته می‌شود. پیچیدگی زمانی الگوریتم، O(Elog(V)) است. خروجی این الگوریتم، درختی با حداقل مجموع وزن است که دارای همان تعداد رأس (N) و N-1 یال در شبکه بازار است.

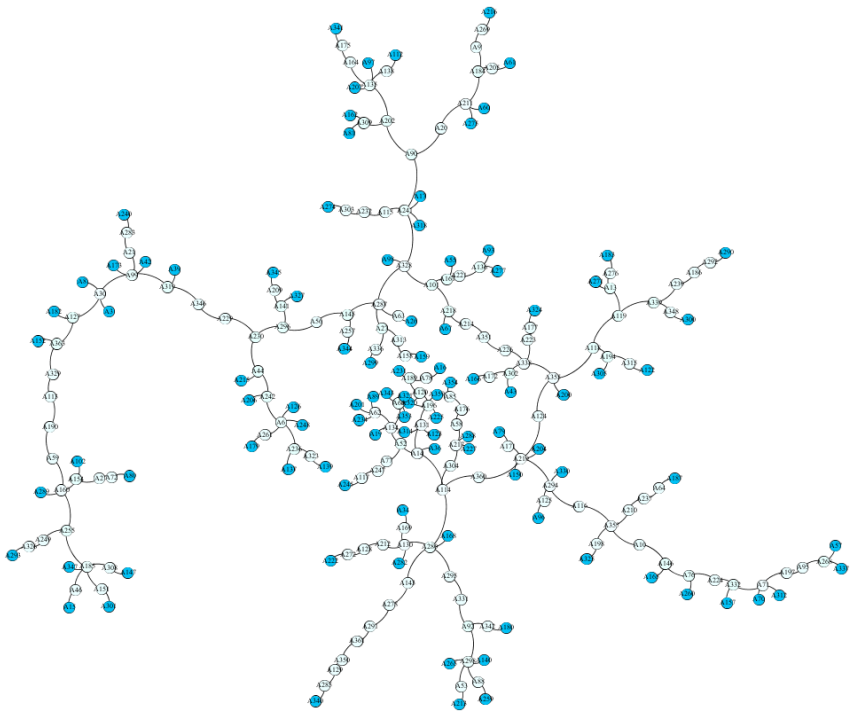
۴- یافته‌های پژوهش

به منظور ایجاد شبکه بازار سهام، از قیمت پایانی تعدیلی روزانه سهام معامله شده طی دوره زمانی ۱۳۹۲/۰۱/۰۷ تا ۱۳۹۸/۱۲/۲۸ برای محاسبه بازده استفاده شده و سپس همبستگی بازده سهام‌ها محاسبه گردیده است. مقدار آستانه بهینه بر اساس متدولوژی معرفی شده توسط زو و همکاران (۲۰۱۸) و آزمایش انتخاب‌های متوالی θ برابر ۰/۴ محاسبه شده و شبکه حاصل با ۷۲۰۰ یال و ۲۴۶ گره تشکیل می‌گردد.

به منظور دست‌یابی به اطلاعات ضروری و قابل اتکا، از مدل‌سازی فیلترینگ شبکه بر اساس MST استفاده می‌شود. الگوریتم کروسکال این مزیت را دارد که با پیچیدگی کمتری نسبت به سایر الگوریتم‌های مدرن، مسأله یافتن MST در شبکه بازار سهام ایران را تحلیل می‌کند. یال‌ها، اطلاعات یا نوسانات را در امتداد شاخه‌های MST عبور می‌دهند. بر اساس وندولا و همکاران[1] (۲۰۰۱) انتهاهای معلق در MST به میزان کمتری تحت تأثیر نوسانات بازار قرار می‌گیرند و با سهام‌های غالب و یا مرکزی همبستگی کمتری دارند. از سوی دیگر بر اساس بوگینسکی و همکاران (۲۰۰۶) گروهی از سهام که دارای حداقل همبستگی و یا همبستگی منفی هستند، می‌توانند مجموعه‌ای مستقل را تشکیل دهند که به خوبی متنوع‌سازی شده هستند. با توجه به این ایده‌ها، در این تحقیق با رویکرد انتهاهای معلق[2] در MST، سهام‌های انتخابی گزینش می‌شوند تا سبدی شامل سهام‌های با همبستگی کمتر و با تنوع بیشتر ارائه شود.

با مدل‌سازی فیلترینگ شبکه بازار سهام بر اساس MST، تعداد روابط بین سهام‌ها به ۲۴۵ یال کاهش می‌یابد. این بدین معنا است که از میان تمام یال‌های موجود در شبکه بازار، ۲۴۵ یال باقی مانده مهم‌ترین روابط را شکل می‌دهند و توانایی نمایش کل بازار را به صورت جامع دارند. همان‌طور که در شکل ۱ ملاحظه می‌شود تنها یک مسیر بین هر جفت سهام وجود دارد که مسیر طولانی‌تر شامل سهام‌های با همبستگی کمتر است. از این رو، در این تحقیق سهام‌های انتخاب شده مربوط به گره‌هایی با درجه یک هستند که به صورت توپر در شکل ۱ نشان داده شده‌اند.

[1] . Vandewalle et al.
[2] . Dangling ends



شکل ۱- مدل‌سازی فیلترینگ شبکه بازار سهام ایران بر اساس MST

مأخذ: یافته‌های پژوهشگر

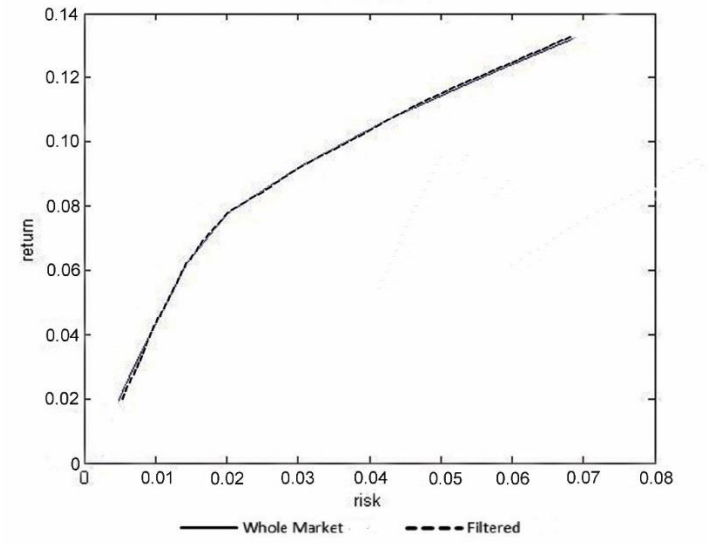
سهم صنایع در این مدل‌سازی، به صورت جدول ۱ است. همان‌طور که مشاهده می‌شود صنایع محصولات شیمیایی و دارویی سهم نسبتاً بیشتری از سایر صنایع در فیلترینگ شبکه بازار سهام بر اساس MST دارند.

جدول ۱- سهم صنایع در مدل‌سازی فیلترینگ شبکه بازار سهام

محصولات شیمیایی	محصولات دارویی	سیمان، آهک و گچ	فلزات اساسی	خودرو	محصولات غذایی به جز قند و شکر	بانک‌ها و موسسات اعتباری	انبوه‌سازی	سرمایه‌گذاری‌ها	کانی‌های غیرفلزی	بیمه	قند و شکر	سایر
۹٪/۱	۷٪/۲	۶٪/۹	۶٪/۷	۶٪/۷	۶٪/۴	۶٪/۱	۵٪/۸	۵٪/۴	۵٪/۱	۴٪/۵	۳٪/۱	۲۶٪/۲

مأخذ: یافته‌های پژوهشگر

به منظور ارزیابی اثربخشی مدل فیلترینگ شبکه بازار سهام لازم است بررسی شود که فیلترسازی بازار سهام بر اساس MST، اطلاعات مهمی را حذف نکرده است. برای این هدف، مرز کارای حاصل از سبد فیلترینگ شبکه بازار سهام با مرز کارای حاصل از سبد کل بازار مقایسه می‌گردد. مرز کارا را می‌توان با منحنی مبادله ریسک-بازده سبد بر مبنای مدل مارکوویتز ترسیم کرد. ریسک در این مطالعه بر اساس معیار انحراف معیار محاسبه شده است. بر اساس مدل مارکوویتز، مدل برنامه‌ریزی ریاضی حداقل‌سازی واریانس سبد به ازای مقدار بازده معین حل می‌گردد و بردار وزن‌های سبد بدین ترتیب تعیین می‌شود. با تغییر مقادیر بازده، منحنی تبادل ریسک-بازده ترسیم می‌گردد که به مرز کارا موسوم است. مقایسه عملکرد ریسک-بازده این دو سبد با دو منحنی مرز کارا در شکل ۲ ارائه شده است.



شکل ۲- مرز کارای حاصل از سبد کل بازار و سبد حاصل از فیلترینگ شبکه بازار سهام بر اساس MST

مأخذ: یافته‌های پژوهشگر

واضح‌ترین نتیجه‌ای که از شکل ۲ و بررسی دو منحنی مرز کارا حاصل می‌شود این است که تفاوتی میان عملکرد ریسک-بازده سبد کل بازار سهام و سبد حاصل از مدل‌سازی فیلترینگ شبکه بازار سهام بر اساس MST وجود ندارد و رفتار فیلترینگ شبکه با حذف یال‌های غیر ضروری، مشابه کل بازار می‌باشد. بدین ترتیب سهام حاصل از فیلترینگ شبکه با کاهش قابل توجهی در اندازه، رفتار کل بازار را تقلید می‌نماید.

برای تأیید این ادعا که فیلترینگ شبکه بازار سهام بر اساس MST، از تنوع مشابهی با کل بازار برخوردار است، از شاخص متنوع‌سازی[1] معرفی شده توسط رودین و مورگان[2] (۲۰۰۶) استفاده می‌شود. این معیار با استفاده از تحلیل مؤلفه‌های اصلی به صورت زیر محاسبه می‌شود:

$$PDI = 2\sum_{k=1}^{n} k\lambda_k - 1 \qquad (۶)$$

که n، تعداد مؤلفه‌های اصلی سبد و λ_k مقادیر ویژه نرمال‌شده و منظم ماتریس همبستگی است. این معیار در بازه $1 \leq PDI \leq n$ قرار می‌گیرد و زمانی که همه مقادیر ویژه مساوی هستند، $PDI = n$ خواهد بود (حالت آرمانی). بیشتر بودن PDI نشان‌دهنده منابع غیر وابسته بوده و این به معنی تنوع‌سازی بیشتر است.

به منظور مقایسه شاخص متنوع‌سازی ارائه شده در دو مجموعه با اندازه متفاوت، از نسبت $\frac{PDI}{n}$ با لحاظ همسان‌سازی مخرج کسر استفاده می‌شود که $\frac{1}{n} \leq \frac{PDI}{n} \leq 1$. روشن است که هرچه نسبت $\frac{PDI}{n}$ به یک نزدیک‌تر باشد، نشان دهنده متنوع‌سازی بیشتر است. نتایج محاسبه تعدیل‌شده این شاخص برای مجموعه کل بازار سهام و مجموعه حاصل از مدل‌سازی فیلترینگ شبکه بازار سهام بر اساس MST در جدول ۲ ارائه شده است.

جدول ۲- مقایسه متنوع‌سازی کل بازار و شبکه فیلتر شده بازار سهام

	سبد کل بازار	سبد حاصل از فیلترینگ شبکه
نسبت $\frac{PDI}{n}$ تعدیل‌شده	۰/۶۸۹	۰/۶۷۸

مأخذ: یافته‌های پژوهشگر

نتایج حاصل از جدول ۲ حاکی از تنوع‌سازی مشابه کل بازار سهام و شبکه فیلتر شده بازار بر اساس MST است. بر اساس نتایج حاصل از مقایسه متنوع‌سازی کل بازار سهام و شبکه فیلتر شده، مدل‌سازی فیلترینگ شبکه بازار سهام بر اساس MST، مجموعه‌ای کاهش یافته از بازار سهام را ارائه می‌دهد که با کاهش قابل توجه در اندازه، از عملکرد و تنوع‌سازی مشابهی برخوردار است.

۵- نتیجه گیری و بحث

داده‌کاوی بازار سهام یکی از مهم‌ترین ابزارهایی است که قادر به کشف الگوهای پنهان در بازار سهام می‌باشد. بازار سهام، پایگاه‌های داده بزرگی را تولید می‌نماید که استفاده از روش‌های داده‌کاوی، بررسی این اطلاعات پیچیده را تسهیل می‌کند. از این رو، روش‌های داده‌کاوی به عنوان روشی کارآمد برای خلاصه کردن و تجسم داده‌های بازار سهام از اهمیت خاصی در بازارهای مالی برخوردار است.

[1] . Portfolio Diversification Index
[2] . Rudin & Morgan

در این مقاله، از داده‌های موجود در بازار سهام ایران برای ساخت شبکه استفاده شد و بر مبنای روش آستانه، شبکه بازار سهام ایران شکل گرفت و سپس فیلترینگ شبکه بر اساس می‌نیمم درخت فراگیر صورت گرفت. بر اساس این رویکرد، یال‌های اضافی در شبکه حذف شده و ارتباطات مهم و قابل اتکا در شبکه باقی می‌مانند. به منظور بررسی اثربخشی فیلترینگ شبکه بازار سهام، مجموعه کل بازار سهام و مجموعه حاصل از فیلترینگ شبکه بازار سهام مقایسه گردید. نتایج حاصل از مقایسه این دو مجموعه نشان داد مدل‌سازی فیلترینگ شبکه بازار سهام ایران بر اساس MST می‌تواند زیرمجموعه‌ای از بازار سهام را تشکیل دهد که عملکرد ریسک-بازده کل بازار را با کاهش قابل توجهی در اندازه دنبال نماید و از درجه تنوع‌سازی مشابهی با کل بازار برخوردار باشد.

بنابراین فیلترینگ شبکه مبتنی بر MST در این مقاله، زیرمجموعه‌ای از بازار سهام را با کاهش قابل توجهی در اندازه ارائه می‌دهد که می‌تواند به عنوان نماینده‌ای مناسب، نماگر ویژگی‌های کل بازار سهام باشد. بدین ترتیب، این مقاله با استفاده از روش‌های داده‌کاوی بازار سهام، جنبه‌های ویژه‌ای از بازار سهام را ضمن کاهش ابعاد ارائه می‌دهد؛ به طوری که ضمن دارا بودن رفتار مشابه کل بازار، ریسک تحلیل با داده‌های بزرگ بازار سهام را نیز کاهش می‌دهد. لازم به ذکر است که در جهت توسعه این پژوهش و عمق بخشیدن به نتایج این مطالعه، سازمان‌ها و نهادهای مربوطه می‌توانند نتایج این مطالعه را در طول سال‌های بیشتری مورد بررسی قرار دهند که با توجه به دشواری تجزیه و تحلیل داده‌های بزرگ، از حوصله این پژوهش خارج است.

فهرست منابع

۱) اسماعیل‌پورمقدم، هادی؛ تیمور محمدی؛ محمد فقهی‌کاشانی و عباس شاکری. (۱۳۹۸). ارائه شاخصی جدید برای انعکاس رفتار بازار سهام با استفاده از رویکرد تحلیل شبکه‌های پیچیده. فصلنامه اقتصاد مالی، ۴۶: ۳۹-۲۵.

۲) اسماعیل‌پورمقدم، هادی؛ تیمور محمدی؛ محمد فقهی‌کاشانی و عباس شاکری. (۱۳۹۸). رشد بخشی و مرکزیت در بازار سهام ایران: رویکرد تحلیل شبکه‌های پیچیده. فصلنامه پژوهش‌ها و سیاست‌های اقتصادی، ۹۰: ۳۱۳-۳۴۱.

۳) خطیب سمنانی، محمدعلی؛ سمیه ایزدی و سیدرضا خادمی (۱۳۹۶). مزیت رقابتی داده‌کاوی در بانکداری الکترونیکی در برابر نااطمینانی‌های اقتصادی. فصلنامه اقتصاد مالی، ۳۹: ۱۴۵-۱۲۵.

۴) ولایتی، محمد؛ فرهاد حسین‌زاده لطفی؛ محمدرضا شهریاری و فریدون رهنمای رودپشتی. (۱۳۹۶). رویکرد داده‌کاوی در بخش‌بندی بازار مشتریان به منظور اتخاذ استراتژی‌های کارا (مطالعه موردی صنعت مخابرات). فصلنامه اقتصاد مالی، ۴۱: ۲۶۶-۲۴۳.

5) Allen, F. & Babus, A. (2009). Networks in finance. In The Network Challenge: Strategy, Profit, and Risk in an Interlinked World; Kleindorfer, P.R., Wind, Y., Eds.; Pearson Education, Inc.: Upper Saddle River, NJ, USA, pp. 367–382.

6) Bazlamacci, C. F. & Hindi, K. S. (2001). Minimum - weight spanning tree algorithms: A survey and empirical study. Computers & Operations Research, 28: 767-785.

7) Birch J., Pantelous A. A. & Soramake, K. (2015). Analysis of Correlation Based Networks Representing DAX 30 Stock Price Returns. Computational Economics, 47: 501-525.
8) Boginski, V., Butenko, S.& Pardalos, P. M. (2006). Mining market data: A network approach. Computers & Operations Research, 33: 3171-3184.
9) Brida, J. G., Matesanz, D. & Seijas, M. N. (2016). Network analysis of returns and volume trading in stock markets: The Euro Stoxx case. Physica A, 55: 751–764.
10) Caraiani, P. (2012). Characterizing emerging European stock markets through complex networks: From local properties to self-similar characteristics. Physica A, 391: 3629–3637.
11) Coletti, P. (2016). Comparing minimum spanning trees of the Italian stock market using returns and volumes. Physica A, 463: 246–261.
12) De Carvalho, P. J. C., & Gupta, A. (2018). A network approach to unravel asset price comovement using minimal dependence structure. Journal of Banking & Finance, 91, 119-132.
13) Dimitrios, K. & Vasileios, O. (2015). A Network Analysis of the Greek Stock Market. Procedia Economics and Finance, 33: 340-349.
14) Eberhard, J., Lavin, J. F. & Montecinos-Pearce, A. (2017). A Network-Based Dynamic Analysis in an Equity Stock Market. Complexity, 17: 1-16.
15) Elton, E.J. & Gruber, M.J. (1995). Modern Portfolio Theory and Investment Analysis, New York, Wiley.
16) Esmaeilpour Moghadam, H., Mohammadi, T., Feghhi Kashani, T. & Shakeri, A. (2019). Complex networks analysis in Iran stock market: The application of centrality, Physica A, 121800.
17) Eom, C., Oh, G., Jung, W. S., Jeong, H. & Kim, S. (2009). Topological properties of stock networks based on minimal spanning tree and random matrix theory in financial time series. Physica A, 388: 900 – 906.
18) Fiedor, P. (2014). Networks in financial markets based on the mutual information rate. Physical Review, 89(5): 14-22.
19) George, S. & Changat, M. (2017). Network approach for stock market data mining and portfolio analysis, International Conference on Networks & Advances in Computational Technologies (NetACT), Thiruvanthapuram, 251-256.
20) Jallo, D., Budai, D. Boginski, V. Goldengorin, B. & Pardalos, P.M. (2013). Network-based representation of stock market dynamics: an application to American and Swedish stock markets models. Algorithms Technol. Netw. Anal: 32, 93–106.
21) Jo, S. K., Kim, M. J., Lim, K., & Kim, S. Y. (2018). Correlation analysis of the Korean stock market: Revisited to consider the influence of foreign exchange rate. Physica A: Statistical Mechanics and its Applications, 491, 852-868.
22) Kheyrkhah, A., Rahnamay Roodposhti, F. & Aliafsar Kazemi, M. (2016). Using the Theory of Network in Finance. International Journal of Finance and Managerial Accounting: 2, 9-23.
23) Kito, T., Ueda, K. (2014). The implications of automobile parts supply network structures: A complex network approach. CIRP Annals-Manufacturing Technology, 63(1), 393-396.
24) Kruskal, J. B. (1956). On the Shortest Spanning Subtree of a Graph and the Traveling Salesman Problem. Proceedings of the American Mathematical Society, 19: 56-65.
25) Laloux, L., Cizeau, P., Cotters, M. & Bouchaud. J. (2000). Random matrix theory and financial correlations. Mathematical models and methods in applied sciences, 3: 1-7.
26) Long, W., Guan, L., Shen, J., Song, L. & Cui, L. (2017). A complex network for studying the transmission mechanisms in stock market. Physica A: 484, 345–357.
27) Majapa, M. & Gossel, S. J. (2016). Topology of the South African stock market network across the 2008 financial crisis. Physica A, 445: 35–47.
28) Nier, E., Yang, J., Yorulmazer, T. & Alentorn, A. (2008). Network models and financial stability. Working Paper No. 346, Bank of England.

29) Pecora, N. & Spelta, A. (2015). Shareholding relationships in the Euro Area banking market: A network perspective. Physica A: Statistical Mechanics and its Applications, 43: 1-12.
30) Rudin, A. M., & Morgan, J. S. (2006). A portfolio diversification index. The Journal of Portfolio Management, 32(2): 81–89.
31) Sharma, K., Shah, S., Chakrabarti A.S., & Chakraborti, A. (2017) Sectoral Co-movements in the Indian Stock Market: A Mesoscopic Network Analysis. In: Aruka Y., Kirman, A. (eds) Economic Foundations for Social Complexity Science. Evolutionary Economics and Social Complexity Science, Singapore: Springer.
32) Shi, Y., Li, L., Wang, Y., Chen, J., & Stanley, H. E. (2019). A study of Chinese regional hierarchical structure based on surnames. Physica A: Statistical Mechanics and its Applications, 518, 169-176.
33) Song, J. W., Ko, B., & Chang, W. (2018). Analyzing systemic risk using non-linear marginal expected shortfall and its minimum spanning tree. Physica A: Statistical Mechanics and its Applications, 491, 289-304.
34) Tong, C., Chen, J., & Buckle, M. J. (2018). A network visualization approach and global stock market integration. International Journal of Finance & Economics.
35) Tse, C. K., Liu, J. & Lau, F. C. M. (2010). A network perspective of the stock market. Journal of Empirical Finance, 17: 659–667.
36) Tumminello, M., Lillo, F. & Mantegna, R. N. (2010). Correlation, hierarchies, and networks in financial markets. Journal of Economic Behavior & Organization, 75: 40 – 58.
37) Vandewalle, N., Brisbois, F. & Tordoir, X. (2001). Non - random topology of stock markets. Quantitative Finance, 1: 372 - 374.
38) Xu, X. J., Wang, K., Zhu, L., & Zhang, L. J. (2018). Efficient construction of threshold networks of stock markets. Physica A: Statistical Mechanics and its Applications, 509, 1080-1086.
39) Yang, C., Chen, Y., Niu, L. & Li, Q. (2014). Cointegration analysis and influence rank-A network approach to global stock markets. Physica A, 400: 168–185.
40) Zhang, J., Zhou, H. & Jiang, L. (2010). Network topologies of Shanghai stock index. Physics Procedia, 3(5):1733–1740.
41) Zhang, Y., Cao, X., He, F. & Zhang, W. (2017). Network topology analysis approach on China's QFII stock investment behavior. Physica A: Statistical Mechanics and its Applications, 473: 77-88.
42) Zhao, L., Li, W.& Cai, X. (2016). Structure and dynamics of stock market in times of crisis. Physics Letters A, 380: 654–666.
43) Zhong, T., peng, Q. Wang, X. & Zhang, J. (2016). Novel indexes based on network structure to indicate financial market. Physica A, 443: 583–594.

Financial Economics
Vol. (17) Issue (62) April 2023

Abstract

https://doi.org/10.30495/fed.2023.1934542.2455

Data mining of Iranian stock market by modeling complex network filtering based on MST

Hadi Esmaeilpour Moghadam [1]

Received: 26 / December / 2022 Accepted: 26 / February / 2023

Abstract

One of the most important problems in modern finance is finding efficient ways to summarize and visualize stock market data. Modeling the filtering of complex networks in the stock market allows this to be achieved by reducing the market size, obtaining reliable information with less disturbance. Since stock price changes are not independent of each other, the study of the correlation between stock price changes provides a better understanding of market performance for investors. Stock market analysis based on complex networks allows studying the correlation of stock prices. In this paper, using the stock market data in the Tehran Stock Exchange, the Iranian stock market network is created by the threshold method, and then the network filtering is based on MST. The results show that the filtration modeling of Iran's stock market network based on the MST can form a subset of the stock market that follows the performance of the entire market with a significant reduction in size and has a similar degree of diversification with the entire market. These analyzes provide a more in-depth insight into the structure of the stock market while reducing the size.

Keywords: Stock market filtering, complex network, MST, diversification
JEL Classification: D53, G11, G20

[1] Economics, Ferdowsi University, Mashhad, Iran. (corresponding author), esmaeilpour@um.ac.ir

This work is licensed under a Creative Commons Attribution 4.0 International License