



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

هادی اسماعیل پورمقدم^۱

تاریخ پذیرش: ۱۴۰۱/۱۲/۰۷

تاریخ دریافت: ۱۴۰۱/۱۰/۰۵

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. به علاوه، 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 می‌نیم درخت فراگیر (MST) صورت می‌گیرد. نتایج نشان می‌دهد مدل‌سازی فیلترینگ شبکه بازار سهام ایران بر اساس می‌نیم 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، می‌نیم درخت فراگیر، متنوع‌سازی.

طبقه بندی JEL: D53, G11, G20

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



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 (تسی 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 (تسی 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 (پکروا 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 تأثیر 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

^۱ . Boginski et al.

^۲ . Dimitrios & Vasileios

^۳ . Tse et al.

^۴ . Nier et al.

^۵ . Jallo et al.

^۶ . Pecora & Spelta

^۷ . Kheyrkhah et al.

کارآمد برای شبکه بازار سهام است که یال های^۱ اضافی از شبکه بازار سهام حذف شود. به طور کلی، فرایندهای فیلترسازی که در ادبیات شبکه استفاده می شود، اغلب با فیلتر اتصال ها یا یال ها به حذف اختلال از ماتریس همبستگی شبکه بازار اقدام می کنند تا ماتریس همبستگی را قابل اتکاتر برای بهینه سازی سید نمایند. روش می نیم درخت فراگیر (MST^۲) یکی از معمول ترین و شناخته شده ترین این روش ها است که درختی پوشا از گرافی متصل بدون جهت را شکل می دهد؛ به طوری که جمع اوزان یال های انتخاب شده حداقل باشد. از این رو، این مطالعه در تلاش است با ایجاد شبکه ای کامل از ارتباطات مبتنی بر همبستگی برای سهام های بازار بورس ایران، اطلاعات مربوط به ساختار داخلی بازار سهام را منعکس نماید. سپس با مدل سازی فیلترینگ شبکه بر اساس MST، یال های اضافی را حذف نموده و زیرمجموعه ای 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 تنوع سازی مشابهی با کل بازار برخوردار باشد. مزیت این رویکرد، در کاهش هزینه و ریسک بررسی داده های بزرگ بازار با رصد گروه نسبتاً اندکی از سهام به جای کل بازار است. دیگر مزیت مدل سازی فیلترینگ شبکه بر اساس MST این است که با حذف یال های اضافی، امکان درک شهودی آسان تر کل شبکه فراهم می آید که در این تحقیق نیز بدان پرداخته می شود. این نوع نمایش شهودی، راهی بصری برای تشخیص عملکرد سهام و رفتار مشابه آن ها ارائه می دهد. ادامه این مقاله به شرح زیر می باشد: بخش دو، ادبیات موضوع و مرور مطالعات قبلی در زمینه شبکه های پیچیده است. سپس بخش سه، روش تحلیلی را توصیف می کند و در نهایت، به ارائه نتایج و نتیجه گیری پرداخته می شود.

۲. ادبیات نظری و مرور پیشینه پژوهش

تحلیل همبستگی اهمیت قابل توجهی در مدیریت ریسک و تخصیص دارایی دارد (التون و گروبر^۳، ۱۹۹۵). مطالعه ماتریس های همبستگی دارای تاریخچه ای طولانی در امور مالی است و سنگ بنای اصلی نظریه مارکوفیتز در مورد اوراق بهادار است (لالوکس و همکاران^۴، ۲۰۰۰). شبکه بازار سهام مطالعه همبستگی تغییرات قیمتی سهام را فراهم می آورد. استفاده از ابزار شبکه به منظور تجزیه و تحلیل وابستگی های اقتصادی یا به هم پیوستگی های اقتصاد بعد از بحران مالی جهانی ۲۰۰۸-۲۰۰۹ گسترش پیدا کرد که ماهیت شبکه بودن سیستم های بانکی را به عنوان مجموعه ای از شرکت ها که با ارتباط نزدیک با یکدیگر عمل می کنند، آشکار کرد. رویکرد شبکه مربوط به ساختار و تشکیل سیستم گره ها^۵ می باشد. یک فرض بنیادی این است که تماماً ویژگی ها و رفتار یک گره می تواند تنها با توجه به ارتباط آن با بقیه سیستم مورد ارزیابی قرار گیرد. در شبکه های مالی، گره ها معمولاً مؤسسات مالی یا نهادهای مشابه هستند. هنگامی که عوامل با توجه به هزینه و منافع ضمنی، متصل و مرتبط می شوند، این فرآیند اتصال لزوماً بستگی به موقعیت نسبی آن ها در شبکه دارد (آلن و بابوس^۶، ۲۰۰۹).

^۱ . Edge

^۲ . Minimum Spanning Tree

^۳ . Elton & Gruber

^۴ . Laloux et al.

^۵ . Node

^۶ . 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 شما و همکاران^{۱۱} (۲۰۱۷) همبستگی سهام با استفاده از روش آستانه انجام گرفت و شبکه بازار سهام هند بدین مینا 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 سید 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)،

^۱ . Caraiani

^۲ . Yang et al.

^۳ . Coletti

^۴ . Majapa & Gossel

^۵ . Brida et al.

^۶ . Zhong et al.

^۷ . Zhao et al.

^۸ . Zhang et al.

^۹ . Long et al.

^{۱۰} . Eberhard et al.

^{۱۱} . Sharma et al.

^{۱۲} . George & Changat

^{۱۳} . Esmailpour Moghadam et al.

^{۱۴} . Eom et al.

^{۱۵} . De Carvalho & Gupta

^{۱۶} . Jo et al.

تنگ و همکاران^۱ (۲۰۱۸)، شی و همکاران^۲ (۲۰۱۹)، سانگ و همکاران^۳ (۲۰۱۸) پیرامون MST، اهمیت این رویکرد را در شبکه بازار سهام تأیید می نماید. تمینلو و همکاران^۴ (۲۰۱۰) از MST به عنوان ابزاری برای فیلتر ارتباطات بین سهام استفاده کردند و خوشه بندی سلسله مراتبی را برای شبکه بازار سهام اعمال نمودند. فیدر^۵ (۲۰۱۴) با اعمال فیلترینگ MST بر شبکه داده های بازار سهام نزدیک نشان داد این رویکرد منجر به نتایج مفیدی برای گسترش پویایی های بازار سهام می گردد. در مطالعه بیرچ و همکاران^۶ (۲۰۱۵) نیز تأیید گردید که مدل سازی فیلترینگ شبکه بازار سهام موجب ارائه بینشی جدید پیرامون پویایی های اقتصاد می گردد.

۳- روش شناسی پژوهش

تئوری شبکه های پیچیده، برگرفته از تئوری گراف می باشد که به عنوان یک چارچوب نظری برای درک ویژگی های ساختاری شبکه ها توسعه یافته است (کیتو و ادا^۷، ۲۰۱۴). شبکه، یک سه تایی $G = (V, E, f)$ است که در آن V مجموعه ای متناهی از گره ها و $E = \{e_1, e_2, \dots, e_m\}$ مجموعه ای از پیوندها یا یال ها و f نگاشتی است که برخی از عناصر E را به یک جفت از عناصر V مربوط می کند؛ به طوری که اگر $v_i \in V$ و $v_j \in V$ باشد، آن گاه خواهیم داشت: $f: e_p \rightarrow [v_i, v_j]$ و $f: e_q \rightarrow [v_j, v_i]$. اگر $p_i(t)$ قیمت پایانی سهم i در روز t باشد، بازده سهام در روز t ام به صورت ذیل است:

$$r_i(t) = \ln p_i(t) - \ln p_i(t-1) \quad (1)$$

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

$$c_{ij} = \frac{\langle r_i r_j \rangle - \langle r_i \rangle \langle r_j \rangle}{\sqrt{(\langle r_i^2 \rangle - \langle r_i \rangle^2)(\langle r_j^2 \rangle - \langle r_j \rangle^2)}} \quad (2)$$

که r بازده و گروه، 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 می گردد و یک یال بدون جهت بین گره های i و j رسم می شود 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 آن ها معرفی می کنند 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 می شود. بر این اساس، برای هر مقدار $\theta_i = [0,1]$ ، ماتریس همبستگی G_i و شبکه N_i ساخته می شود و تابع ذیل محاسبه می گردد:

$$G_{\theta_i} = \frac{\langle D_C D_N \rangle - \langle D_C \rangle \langle D_N \rangle}{\sigma_{D_C} \sigma_{D_N}} \quad (3)$$

این تابع، سازگاری بین تغییرات در ماتریس همبستگی و شبکه را اندازه گیری می نماید 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 آن ها است که با روش عددی مقدار آستانه بهینه بر اساس ذیل انتخاب می شود:

$$\hat{\theta} = \arg \max_{\theta_i} \{G_{\theta_i}\} \quad (4)$$

هم چنین مطابق با ژنگ و همکاران^۱ (۲۰۱۰)، از قدرمطلق همبستگی به عنوان وزن اختصاص داده شده به یال در شبکه استفاده می گردد. در این صورت، $G = (V, E, W)$ نشان دهنده شبکه سهام است که V مجموعه ای از رأس ها را نشان می دهد، E نشان دهنده یال ها و W وزن یال است. W به شرح زیر تعریف می شود:

$$W = \begin{cases} w_{ij} = |c_{ij}|, & i \neq j \text{ and } |c_{ij}| \geq 0 \\ w_{ij} = 0, & \text{else.} \end{cases} \quad (5)$$

اگر $w_{ij} \neq 0$ ، آن گاه یک یال بین گره های i و j وجود خواهد داشت.

گام دوم پس از ایجاد شبکه بازار سهام، فیلترسازی شبکه با استفاده از می نیمم درخت فراگیر (MST) است. MST ابزاری شناخته شده در ادبیات شبکه برای فیلترسازی یال های اضافی از شبکه است. بدین منظور چنانچه درخت $T(V, E')$ به عنوان زیرمجموعه ای 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 می شود؛ به طوری که $v_i, v_j \in V$ و $e_{ij} \in E$. درخت یکتای T برای گراف G ، MST است اگر و فقط اگر مقدار تابع $\sum_{e \in E} W(e)$ می نیمم باشد.

الگوریتم های حل MST توسط بزلامسی و هندی^۲ (۲۰۰۱) بررسی شده و بر اساس اندازه شبکه و عملکرد زمانی به دو طبقه بندی کلاسیک و مدرن تقسیم شده اند. الگوریتم های مدرن که اخیراً توسعه یافته اند، از روش های جستجوی تصادفی برای حل MST استفاده می کنند که باعث می شود مدل با سرعت بیشتری نسبت به الگوریتم های کلاسیک اجرا شود. از این رو، الگوریتم های مدرن پیچیده تر هستند؛ اما برای مسائل با مقیاس بزرگ کارآمدتر از الگوریتم های کلاسیک هستند. در این تحقیق، از الگوریتم کروسکال^۳ به عنوان الگوریتمی مدرن برای حل MST استفاده می شود. الگوریتم کروسکال، یک الگوریتم حریصانه است که در آن یال ها بر اساس وزن به طور صعودی مرتب می شوند. در ابتدا یال e_1 با حداقل وزن انتخاب شده و سپس به ترتیب یال های دیگر، چنانچه نقطه پایانی

^۱ . Zhang et al.

^۲ . Bazlamacci & Hindi

^۳ . Kruskal

یال در همان درخت باشد، انتخاب می گردند؛ در غیر این صورت، آن یال نادیده گرفته می شود. پیچیدگی زمانی الگوریتم، $O(E \log(V))$ 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- یافته های پژوهش

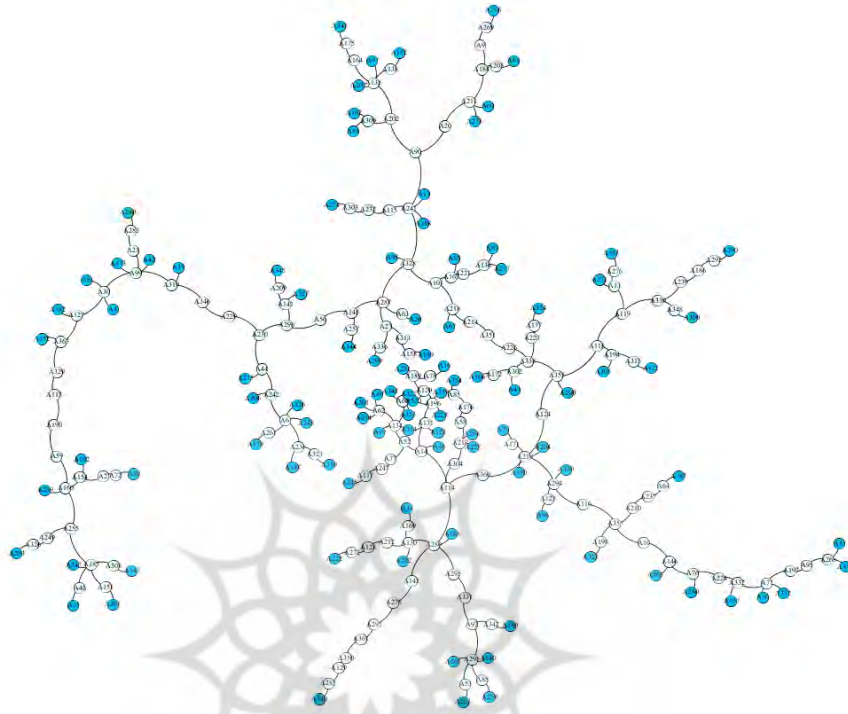
به منظور ایجاد شبکه بازار سهام، از قیمت پایانی تعدیلی روزانه سهام معامله شده طی دوره زمانی ۱۳۹۲/۰۱/۰۷ تا ۱۳۹۸/۱۲/۲۸ برای محاسبه بازده استفاده شده و سپس همبستگی بازده سهامها محاسبه گردیده است. مقدار آستانه بهینه بر اساس متدولوژی معرفی شده توسط زو و همکاران (۲۰۱۸) و آزمایش انتخاب های متوالی θ برابر ۰/۴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

به منظور دستیابی به اطلاعات ضروری و قابل اتکا، از مدل سازی فیلترینگ شبکه بر اساس MST استفاده می شود. الگوریتم کروسکال این مزیت را دارد که با پیچیدگی کمتری نسبت به سایر الگوریتم های مدرن، مسأله یافتن MST در شبکه بازار سهام ایران را تحلیل می کند. یال ها، اطلاعات یا نوسانات را در امتداد شاخه های MST عبور می دهند. بر اساس وندولا و همکاران^۱ (۲۰۰۱) انتهاهای معلق در MST به میزان کمتری تحت تأثیر نوسانات بازار قرار می گیرند و با سهام های 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، می توانند مجموعه ای مستقل را تشکیل دهند که به خوبی متنوع سازی شده هستند. با توجه به این ایده ها، در این تحقیق با رویکرد انتهاهای معلق^۲ در MST، سهام های انتخابی گزینش می شوند تا سبدهی شامل سهام های با همبستگی کمتر و با تنوع بیشتر ارائه شود.

با مدل سازی فیلترینگ شبکه بازار سهام بر اساس MST، تعداد روابط بین سهام ها به ۲۴۵ یال کاهش می یابد. این بدین معنا است که از میان تمام یال های موجود در شبکه بازار، ۲۴۵ یال باقی مانده مهم ترین روابط را شکل می دهند و توانایی نمایش کل بازار را به صورت جامع دارند. همان طور که در شکل ۱ ملاحظه می شود تنها یک مسیر بین هر جفت سهام وجود دارد که مسیر طولانی تر شامل سهام های با همبستگی کمتر است. از این رو، در این تحقیق سهام های انتخاب شده مربوط به گره هایی با درجه یک هستند که به صورت توپر در شکل ۱ نشان داده شده اند.

^۱ . Vandewalle et al.

^۲ . Dangling ends



شکل ۱- مدل سازی فیلترینگ شبکه بازار سهام ایران بر اساس MST

مأخذ: یافته های پژوهشگر

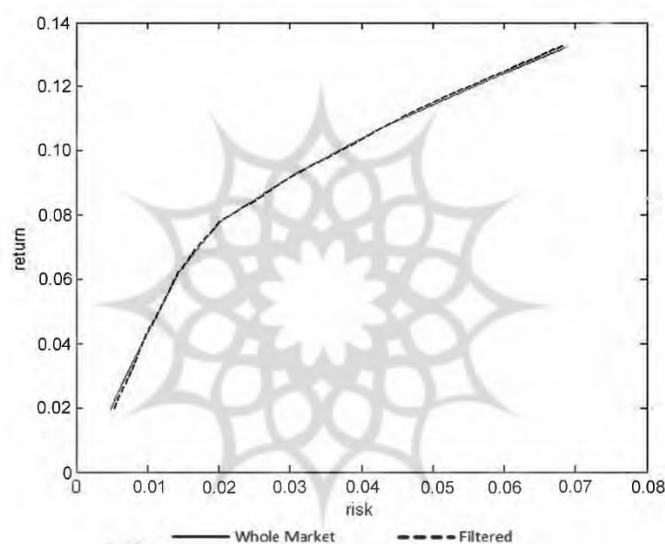
سهام صنایع در این مدل سازی، به صورت جدول ۱ است. همان طور که مشاهده می شود صنایع محصولات شیمیایی و دارویی سهم نسبتاً بیشتری از سایر صنایع در فیلترینگ شبکه بازار سهام بر اساس MST دارند.

جدول ۱- سهم صنایع در مدل سازی فیلترینگ شبکه بازار سهام

محصولات شیمیایی	محصولات دارویی	سیمان، آهک و گچ	فلزات اساسی	خودرو	محصولات غذایی به جز قند و شکر	بانک ها و موسسات اعتباری	انبوه سازی	سرمایه گذاری ها	کانی های غیر فلزی	بیمه	قند و شکر	سایر
۹٪/۱	۷٪/۲	۶٪/۹	۶٪/۷	۶٪/۷	۶٪/۴	۶٪/۱۱	۵٪/۱۸	۵٪/۱۴	۵٪/۱۱	۴٪/۱۵	۳٪/۱۱	۲۶٪/۱۲

مأخذ: یافته های 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 می گردد. مرز کارا را می توان 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 برنامه ریزی ریاضی حداقل سازی واریانس سبد به ازای مقدار بازده معین حل می گردد و بردار وزن های سبد بدین ترتیب تعیین می شود. با تغییر مقادیر بازده، منحنی تبادل ریسک-بازده ترسیم می گردد 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: یافته های پژوهشگر

واضح ترین نتیجه ای که از شکل ۲ و بررسی دو منحنی مرز کارا حاصل می شود این است که تفاوتی میان عملکرد ریسک-بازده سبد کل بازار سهام و سبد حاصل از مدل سازی فیلترینگ شبکه بازار سهام بر اساس MST وجود ندارد و رفتار فیلترینگ شبکه با حذف یال های غیر ضروری، مشابه کل بازار می باشد. بدین ترتیب سهام حاصل از فیلترینگ شبکه با کاهش قابل توجهی در اندازه، رفتار کل بازار را تقلید می نماید.

برای تأیید این ادعا که فیلترینگ شبکه بازار سهام بر اساس MST، از تنوع مشابهی با کل بازار برخوردار است، از شاخص متنوع سازی^۱ معرفی شده توسط رودین و مورگان^۲ (۲۰۰۶) استفاده می شود. این معیار با استفاده از تحلیل مؤلفه های اصلی به صورت زیر محاسبه می شود:

$$PDI = 2 \sum_{k=1}^n k \lambda_k - 1$$

(۶)

که n ، تعداد مؤلفه های اصلی سبد و λ_k مقادیر ویژه نرمال شده و منظم ماتریس همبستگی است. این معیار در بازه $1 \leq PDI \leq n$ قرار می گیرد و زمانی که همه مقادیر ویژه مساوی هستند، $PDI = n$ خواهد بود (حالت آرمانی). بیشتر بودن PDI نشان دهنده منابع غیر وابسته بوده و این به معنی تنوع سازی بیشتر است. به منظور مقایسه شاخص متنوع سازی ارائه شده در دو مجموعه با اندازه متفاوت، از نسبت $\frac{PDI}{n}$ با لحاظ همسان سازی مخرج کسر استفاده می شود که $1 \leq \frac{PDI}{n} \leq 1$. روشن است که هرچه نسبت $\frac{PDI}{n}$ به یک نزدیک تر باشد، نشان دهنده متنوع سازی بیشتر است. نتایج محاسبه تعدیل شده این شاخص برای مجموعه کل بازار سهام و مجموعه حاصل از مدل سازی فیلترینگ شبکه بازار سهام بر اساس MST در جدول ۲ ارائه شده است.

جدول ۲- مقایسه متنوع سازی کل بازار و شبکه فیلتر شده بازار سهام

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

مأخذ: یافته های پژوهشگر

نتایج حاصل از جدول ۲ حاکی از تنوع سازی مشابه کل بازار سهام و شبکه فیلتر شده بازار بر اساس MST است. بر اساس نتایج حاصل از مقایسه متنوع سازی کل بازار سهام و شبکه فیلتر شده، مدل سازی فیلترینگ شبکه بازار سهام بر اساس MST، مجموعه ای کاهش یافته از بازار سهام را ارائه می دهد که با کاهش قابل توجه در اندازه، از عملکرد و تنوع سازی مشابهی برخوردار است.

۵- نتیجه گیری و بحث

داده کاوی بازار سهام یکی از مهم ترین ابزارهایی است که قادر به کشف الگوهای پنهان در بازار سهام می باشد. بازار سهام، پایگاه های داده بزرگی را تولید می نماید که استفاده از روش های داده کاوی، بررسی این اطلاعات پیچیده را تسهیل می کند. از این رو، روش های داده کاوی به عنوان روشی کارآمد برای خلاصه کردن و تجسم داده های بازار سهام از اهمیت خاصی در بازارهای مالی برخوردار است.

^۱ . Portfolio Diversification Index

^۲ . Rudin & Morgan

در این مقاله، از داده های 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 می نیمم درخت فراگیر صورت گرفت. بر اساس این رویکرد، یال های اضافی در شبکه حذف شده و ارتباطات مهم و قابل اتکا در شبکه باقی می ماند. 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 مدل سازی فیلترینگ شبکه بازار سهام ایران بر اساس MST می تواند زیرمجموعه ای از بازار سهام را تشکیل دهد که عملکرد ریسک-بازده کل بازار را با کاهش قابل توجهی در اندازه دنبال نماید و از درجه تنوع سازی مشابهی با کل بازار برخوردار باشد.

بنابراین فیلترینگ شبکه مبتنی بر MST در این مقاله، زیرمجموعه ای از بازار سهام را با کاهش قابل توجهی در اندازه ارائه می دهد که می تواند به عنوان نماینده ای مناسب، نماگر ویژگی های کل بازار سهام باشد. بدین ترتیب، این مقاله با استفاده از روش های داده کاوی بازار سهام، جنبه های ویژه ای از بازار سهام را ضمن کاهش ابعاد ارائه می دهد؛ به طوری که ضمن دارا بودن رفتار مشابه کل بازار، ریسک تحلیل با داده های بزرگ بازار سهام را نیز کاهش می دهد. لازم به ذکر است که در جهت توسعه این پژوهش و عمق بخشیدن به نتایج این مطالعه، سازمان ها و نهادهای مربوطه می توانند نتایج این مطالعه را در طول سال های بیشتری مورد بررسی قرار دهند که با توجه به دشواری تجزیه و تحلیل داده های بزرگ، از حوصله این پژوهش خارج است.

فهرست منابع

- 1) اسماعیل پورمقدم، هادی؛ تیمور محمدی؛ محمد فقهی کاشانی و عباس شاکری. (۱۳۹۸). ارائه شاخصی جدید برای انعکاس رفتار بازار سهام با استفاده از رویکرد تحلیل شبکه های پیچیده. فصلنامه اقتصاد مالی، ۴۶: ۳۹-۲۵.
- 2) اسماعیل پورمقدم، هادی؛ تیمور محمدی؛ محمد فقهی کاشانی و عباس شاکری. (۱۳۹۸). رشد بخشی و مرکزیت در بازار سهام ایران: رویکرد تحلیل شبکه های پیچیده. فصلنامه پژوهش ها و سیاست های اقتصادی، ۹۰: ۳۱۳-۳۴۱.
- 3) خطیب سمنانی، محمدعلی؛ سمیه ایزدی و سیدرضا خادمی (۱۳۹۶). مزیت رقابتی داده کاوی در بانکداری الکترونیکی در برابر نااطمینانی های اقتصادی. فصلنامه اقتصاد مالی، ۳۹: ۱۴۵-۱۲۵.
- 4) ولایتی، محمد؛ فرهاد حسین زاده لطفی؛ محمدرضا شهریاری و فریدون رهنمای رودپشتی. (۱۳۹۶). رویکرد داده کاوی در بخش بندی بازار مشتریان به منظور اتخاذ استراتژی های 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) Caraianni, 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.

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 Esmailpour Moghadam¹

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