ارزیابی کارایی روش بازاریابی ویروسی ترکیبی با روش خوشه‌بندی شبکه‌ای و مقایسه نتایج

فریدون اوحدی*
دانشگاه علوم تحقیقات، تهران، ایران
fohadi31@kiau.ac.ir

مهرنوش محمدی
دانشگاه آزاد اسلامی، کرج، ایران
mehr_1984@yahoo.com

محمدجعفر تارخ
دانشگاه صنعتی خواجه‌نصیرالدین طوسی، تهران، ایران
mjtarokh@kntu.ac.ir

تاریخ دریافت: ۱۳۹۷/۰۸/۰۶
تاریخ اصلاحات: ۱۳۹۷/۱۲/۱۴
تاریخ پذیرش: ۱۳۹۷/۱۲/۱۸

چکیده

در یک بازار رقابتی، درک نیاز مشتریان و تبلیغات مؤثر، از مهم‌ترین فاکتورهای بقا به شمار می‌آید. گسترش اینترنت و شبکه‌های مجازی، فرصت مناسبی برای شرکت‌ها جهت تبلیغات فراهم کرده است و لذا مطالعه روش‌ها و مدل‌های بازاریابی الکترونیک اهمیت زیادی پیدا می‌کند. یکی از جدیدترین این روش‌ها، بازاریابی ویروسی است که بر تبلیغات دهان به دهان استوار بوده و قدرت بسیار زیادی دارد. بازاریابی ویروسی بر این اصل متکی است که در هر شبکه اجتماعی، تعدادی از کاربران قدرت و تأثیرگذاری بالایی روی دیگران دارند و لذا با شناسایی آن‌ها و ایجاد پیام‌های تبلیغاتی خوب، می‌توان از طریق آن‌ها به بازاریابی مؤثر پرداخت. بنابراین، شناسایی کاربران با اهمیت، مهم‌ترین فعالیت در بازاریابی ویروسی به حساب می‌آید. در این راستا، تحقیقات مختلفی با استفاده از انواع روش‌های مبتنی بر گراف و مبتنی بر انتشار، به شناسایی کاربران پرداخته‌اند. در این پژوهش، از قابلیت‌های هر دو روش استفاده شده و با به‌کارگیری معیار مرکزیت نیمه محلی از روش‌های مبتنی بر گراف و مدل خوشه‌بندی مارکوف از روش‌های مبتنی بر انتشار، یک مدل ترکیبی جدید جهت خوشه‌بندی کاربران و شناسایی کاربران کلیدی، ارائه گردیده است. نتایج به‌دست آمده، نشان‌دهنده همبستگی بالاتر روش پیشنهادی با معیار SIR و در نتیجه کارایی بالاتر آن از سایر روش‌های به‌کار گرفته شده در تحقیق است.

واژگان کلیدی

بازاریابی ویروسی؛ مرکزیت؛ گراف؛ شبکه اجتماعی؛ خوشه‌بندی.

۱- مقدمه

پیشرفت فناوری ارتباطات و گسترش زیرساخت‌های آن، باعث ایجاد ارتباطات گسترده میان شبکه‌های مختلف اجتماعی شده است. افرادی که در مکان‌های متمایز جغرافیایی قرار دارند می‌توانند از طریق شبکه‌های مختلف اجتماعی با یکدیگر ارتباط برقرار کنند. این شبکه‌ها را می‌توان به‌عنوان گراف‌هایی طراحی کرد که مردم آن را تشکیل می‌دهند. در این شبکه‌ها، افرادی که منافع مشترک دارند، جوامع را تشکیل می‌دهند. اعضای این جوامع می‌توانند به‌طور مؤثر اطلاعات را به اشتراک بگذارند و با یکدیگر ارتباط برقرار کنند. انسان‌ها تمایل طبیعی برای تصمیم‌گیری‌های دیگران دارند. در یک شبکه اجتماعی نیز، برخی از افراد به‌عنوان عوامل کلیدی عمل می‌کنند، به‌طوری‌که آن‌ها در معرض نفوذ دیگران قرار می‌گیرند [۶]. چنین فضایی، راه را برای ترویج سریع یک اندیشه، باز کرده است. این چنین است که گروه‌ها و کمپین‌های مختلف و با اهداف متفاوت هر روزه شکل گرفته، گسترش پیداکرده و بسته به قدرت آن، رنگ می‌بازند. در چنین محیطی، علایق مشترکی بین کاربران به سرعت شکل گرفته و البته منجر به رفتارهای متفاوتی می‌گردد. این علایق، می‌تواند شامل گرایش به یک عقیده یا رفتار و یا تحریم و گوشه‌گیرکردن یک سازمان و یا شخص گردد [۱].

نکته مهم و قابل تأمل در این شبکه‌های اجتماعی، این است که در هر شبکه از کاربران، تعدادی از آن‌ها تأثیر به‌سزایی بر دیگران داشته و در واقع نقش کلیدی در شکل‌گیری یک جریان و تبلیغ آن دارند. این مسأله، باعث علاقه زیاد شرکت‌ها به استفاده از این فرصت جهت تبلیغات و رشد شده است. برای مثال، تبلیغات درخصوص استفاده از کالاهای ایرانی و یا تحریم یک کالای خاص، از جمله این جریان‌ها است که باعث گرایش یک اجتماع به استفاده از کالای ایرانی و حمایت از رشد شرکت‌های حاضر در این صنعت شده است. بنابراین، این فرصت برای تمامی سازمان‌ها و شرکت‌ها وجود دارد که با شناسایی کاربران کلیدی و آنالیز رفتاری این شبکه‌ها، به ترغیب جامعه‌ای از کاربران جهت استفاده از کالاهای آنان بپردازند [۲].

بازاریابی ویروسی، تکنیکی مبتنی بر استفاده از توان تبلیغاتی دهان به دهان در جامعه است که با گسترش شبکه‌های اجتماعی، طرفداران زیادی پیدا کرده است. این تکنیک، در واقع راهی برای شناسایی کاربران کلیدی و جلب اعتماد این کاربران است تا به این وسیله، یک محصول از طریق چند معرف شاخص به میلیاردها نفر در شبکه‌های اجتماعی

* نویسنده مسئول

معرفی شوند و سایر کاربران حاضر در جامعه نیز به سمت یک سازمان یا یک کالا گرایش پیدا کرده و این مسأله باعث رشد و افزایش فروش شرکت‌ها گردد [۷]. بازاریابی ویروسی، شبکه‌های اجتماعی را به صورت یک گراف بزرگ تصویر می‌کند که در آن، گره‌های گراف نشان‌دهنده کاربران و یال‌های گراف، نشان‌دهنده سطح ارتباطی گره‌های مختلف با یکدیگر است. برای مثال، تعداد کامنت‌های داده شده در مورد نظرات یک فرد، نشان‌دهنده میزان گرایش جامعه‌ای از کاربران به آن فرد و در واقع سطح اعتماد آن‌ها به آن فرد است و هرچه این تعداد بالاتر باشد؛ نشان‌دهنده کلیدی‌تربودن نقش کاربر مذکور در شبکه است. بنابراین، یال‌های این گراف نشان‌دهنده سطح ارتباطی جهت‌دار کاربران با یکدیگر است و لذا گراف تشکیل‌شده در مدل‌های توسعه‌یافته در بازاریابی ویروسی، معمولاً به صورت جهت‌دار نیز در نظر گرفته می‌شود [۳ و ۴].

نکته مشترک چالشی پیش روی تمامی تحقیقات در حوزه مطالعه شبکه‌های اجتماعی و اثرگذاری کاربران مختلف در آن، این است که حجم داده‌ها بسیار زیاد بوده و تحلیل شبکه‌های بزرگ، بسیار دشوار است. بسیاری از روش‌ها، از رهگیری اطلاعات موفق و پیداکردن سرشاخه‌های آن‌ها استفاده می‌کنند؛ این روش‌ها علی‌رغم منطق ساده و کاملی که دارند؛ فرضیه رفتار تصادفی کاربران و حرکت غیرمنطقی شبکه را در نظر نمی‌گیرند [۵]. روش‌های سنجش مرکزیت سراسری نیز، توانایی کار با شبکه‌های بزرگ را ندارند و علی‌رغم دقت بسیار بالایی که دارند؛ نمی‌توانند در عمل مورد استفاده قرار گیرند. در این میان، مرکزیت محلی راهکاری است که اخیراً در تحقیقاتی مورد استفاده قرار می‌گیرد [۲ و ۵]. این معیارها، توانایی کار با حجم بالای داده‌ها را دارند؛ لیکن از دقت مناسب و قابل رقابت با معیارهای مرکزیت سراسری نیستند.

در این میان، خوشه‌بندی دینامیک، روشی است که می‌توان با استفاده از آن، با دقت بالایی به شناسایی وضعیت شبکه پرداخت و کاربران کلیدی را شناسایی کرد. در روش‌های خوشه‌بندی دینامیک، حرکت از هر گره به گره دیگر، با استفاده از یک تابع احتمال صورت می‌گیرد و قاعدتاً، کیفیت نهایی خوشه‌بندی انجام شده، به دقت تابع احتمال و معیار استفاده شده در آن خواهد داشت [۵].

به این ترتیب، در تحقیق حاضر، از ویژگی‌های مرکزیت نیمه محلی برای تشکیل تابع احتمال روش خوشه‌بندی مارکوف استفاده می‌شود. در این راستا، ابتدا معیارهای مرکزیت محلی و مرکزیت نهایی محاسبه شده و با استفاده از آنها مرکزیت نیمه محلی محاسبه می‌شود. در مورد هر گره، احتمال حرکت به گره‌های مجاور، از نسبت مرکزیت نیمه محلی گره مجاور به مجموع مرکزیت نیمه محلی تمامی گره‌های مجاور استفاده می‌شود. با هر بار شبیه‌سازی، تمامی گره‌ها به صورت دینامیک خوشه‌بندی می‌شوند و در نهایت، گره‌هایی که در بیشترین تکرارها در خوشه کاربران کلیدی طبقه‌بندی شده‌اند؛ به‌عنوان کاربران کلیدی انتخاب می‌شوند.

به این ترتیب، سؤال اساسی تحقیق حاضر این است که چگونه می‌توان روشی برای شناسایی گره‌های با ارزش و مهم در یک شبکه اجتماعی به‌کار گرفت که علاوه بر توانایی‌های لازم برای کار با شبکه‌های بزرگ، دقت بالایی در ارزیابی کاربران داشته باشد و پاسخ ارائه‌شده برای آن نیز، روش ترکیبی خوشه‌بندی دینامیک با استفاده از روش مارکوف و معیار مرکزیت نیمه‌محلی به‌عنوان معیار محاسبه تابع احتمال در الگوریتم خوشه‌بندی است.

با در نظرگرفتن اهمیت بازاریابی ویروسی، مدل‌های مختلفی در این زمینه توسعه داده شده است. در این تحقیق، از ترکیب مدل‌های مرکزیت توسعه داده شده در این تکنیک با روش خوشه‌بندی مارکوف، جهت شناسایی کاربران کلیدی استفاده شده است. به این منظور، در ادامه، مدل‌های مهم توسعه داده شده درخصوص بازاریابی ویروسی مرور می‌شود. سپس، معیارهای اندازه‌گیری مرکزیت کاربران معرفی شده و نتایج به‌دست آمده از آن شرح داده می‌شود. در ادامه، روش خوشه‌بندی کاربران براساس امتیازات اخذشده، معرفی و نتایج آن شرح داده می‌شود. در نهایت، نتایج به‌دست آمده مورد بررسی و تحلیل قرار خواهد گرفت.

۲- پیشینه تحقیق

«دراکر» تعریف جامعی از بازاریابی بیان می‌کند که بازاریابی نه تنها صرفاً برای فروش بیشتر کالا، که برای افزایش فروش با تکیه بر آگاهی و شناخت مشتری و منطبق با نیازهای مشتری انجام می‌شود. در طول زمان، ابزارهای مختلفی جهت بازاریابی به‌کار گرفته شد که قاعدتاً اولین آن‌ها تابلوها و اطلاعیه‌ها بوده است. لیکن، اولین پروژه بازاریابی غیرحضوری را می‌توان تله تل دولت فرانسه دانست که در سال ۱۹۸۱ انجام شد [۳]. گسترش شبکه‌های الکترونیکی و اینترنت همراه با زیرساخت‌های آن شامل کامپیوترهای قابل حمل و قدرتمند و مرورگرهای روان و سریع، باعث گسترش شانس شرکتها برای تبلیغ و البته افزایش اثرگذاری تبلیغات به صورت غیرحضوری شده است. یکی از روش‌های با اهمیت تبلیغات غیرحضوری، بازاریابی ویروسی است که توجه زیادی در سال‌های اخیر را از سمت سرمایه‌داران و محققان جلب کرده است [۸].

بازاریابی ویروسی، نوعی انتقال دهان به دهان اطلاعات است که البته در فضای مجازی و به‌خصوص در شبکه‌های مجازی رخ می‌دهد [۹]. در این شبکه‌ها، همواره تعدادی از افراد، دارای نفوذ بیشتری نسبت به سایرین است و این مسأله باعث ایجاد مزیت برای شرکت‌ها از طریق تأثیرگذاری بر این افراد می‌شود؛ در واقع شرکت به جای تبلیغات در سطح وسیع و هزینه بالا، می‌تواند از طریق این افراد به تأثیرگذاری بر شبکه بپردازد [۱۰]. بنابراین، شناسایی افراد مهم در این شبکه‌ها، مهم‌ترین امر در یک بازاریابی ویروسی موفق است و باعث می‌شود که هزینه پیام‌های بازایابی کاهش یافته و سرعت آن افزایش یابد [۱۱].

تاکنون مدل‌های مختلفی برای بازایابی ویروسی پیشنهاد شده است که در ادامه به‌طور مختصر به آن‌ها پرداخته می‌شود.

- ویروسی ارزشی: این نوع بازایابی در مورد محصولات با کیفیت و ارزشمند است که کیفیت محصول باعث توصیه آن می‌شود [۴].
- ویروسی حیله‌ای: در این نوع بازاریابی، جایزه‌ای که به هر فرد به خاطر ایجاد مشتری جدید تعلق می‌گیرد؛ انگیزه اصلی بازاریابی برای افراد است؛ این نوع بازاریابی را بازاریابی شبکه‌ای نیز می‌گویند [۴].
- ویروسی مارپیچی: این نوع بازاریابی بیشتر در مورد خدمات کاربرد داشته و در واقع انتقال تجربیات استفاده از یک محصول توسط کاربران است [۲].
- ویروسی ناخوشایند: در این حالت، بازاریابی منفی رخ داده و تبلیغات در جهت عکس و تجارب منفی در مورد محصول رخ می‌دهد و باعث کاهش فروش شرکت می‌گردد [۱].

برای توصیف موفقیت در بازایابی ویروسی، می‌توان دو قدم کلی برای آن در نظر گرفت. قدم اول، تمرکز شرکت در ساخت و ایجاد بازاریابی دهان به دهان است که در زمینه شبکه‌های مجازی، اشاره به ساخت پیام‌های پرمخاطب دارد که کاربران به نشر آن علاقمند بوده و در گروه‌ها، کانال‌ها و سایر بسترهای ارتباطی دیگر به اشتراک گذاشته شود [۱۲ و ۱۳]. در قدم دوم، شرکت باید مطمئن شود که این پیام‌های ایجاد شده، به‌دست کاربران هدف شناسایی‌شده در شبکه رسیده و آن‌ها این پیام‌ها را بازنشر می‌دهند که البته می‌توان گفت اهمیت این قدم، به اندازه و یا حتی بیشتر از قدم اول است [۱۴]. عوامل مؤثر بر کارایی بازاریابی ویروسی در قدم اول را در برگرفتن نیازهای گروهی، در برداشتن انگیزه‌های خیرخواهانه و شامل‌شدن انگیزه‌های رشد شخصیتی می‌دانند. به عبارت دیگر، محققان در این پژوهش، عوامل مؤثر بر ایجاد یک پیام با پتانسیل بازنشر زیاد را این عوامل ذکر کرده‌اند [۵]. تحقیقات زیادی نیز به بحث و بررسی قدم دوم یعنی شناسایی کاربران تأثیرگذار پرداخته‌اند که به صورت کلی می‌توان آن‌ها را به روش‌های مبتنی بر شبکه و روش‌های مبتنی بر انتشار طبقه‌بندی کرد. روش‌های مختلف سنجش مرکزیت همچون مرکزیت درجه، مرکزیت نزدیکی، مرکزیت میانداری و مرکزیت بردار ویژه در تحقیقات مختلفی به‌کار گرفته شده‌اند که وابسته به شکل شبکه و فواصل بین گره‌ها و سایر ویژگی‌های ساختاری شکل ظاهری شبکه است [۱۵]. در مقابل، مدل‌های مبتنی بر انتشار، گره‌های مهم را به نحوی با شبیه‌سازی شناسایی می‌کنند [۱۶]. در این روش‌ها، با شروع از یک گره، ویروسی در شبکه منتشرشده و براساس تعداد گره‌های آلوده شده در زمانی مشخص و یا تعداد کل گره‌های طی‌شده در مجموع و سایر ویژگی‌های به این شکل، در مورد اهمیت گره‌ها تصمیم‌گیری می‌شود [۱۷ و ۱۸ و ۶].

با گسترش کسب و کار الکترونیکی، رقابت در حوزه شبکه‌های اجتماعی از نظر تبلیغاتی چه در زمینه‌های تجاری و چه در زمینه‌های فرهنگی و سیاسی نیز بالا گرفت. در این میان، تحقیقات زیادی به بررسی چگونگی اثرگذاری افراد خاصی در یک شبکه اجتماعی پرداخته و به نتایج مشترکی رسیدند که طی آن، معلوم شد که تعداد کمی از کاربران (گره‌ها در یک شبکه اجتماعی)، می‌توانند اثرگذاری بالایی بر پخش یک خبر یا اثرگذاری بر یک دیدگاه، داشته باشند [۱۹ و ۲۰]. در پژوهشی که در سال ۲۰۱۶ انجام شده است؛ محققان چگونگی سرایت اطلاعات در شبکه‌های اجتماعی را با چگونگی انتقال یک ویروس مقایسه کرده و به این نتیجه رسیدند که روند انتشار این دو، شباهت بالایی دارند. محققان، با به‌کاربردن یک روش غیرریاضی ابتکاری، از روش‌های شناسایی ویروس‌ها برای شناسایی این جریانات در شبکه‌های اجتماعی و پیداکردن منشاء آن‌ها استفاده کرده‌اند [۲۰].

در دیگر تحقیقات انجام‌شده، به بحث چگونگی استفاده از بازاریابی ویروسی برای جذب مشتری و افزایش فروش پرداخته شده است. بدین ترتیب که با استفاده از مطالعه انتشار کلمات کلیدی مشخص در یک شبکه اجتماعی، سیر به‌کارگیری این کلمات رهگیری شده و با مطالعه سرعت پخش و گستره پخش اطلاعات، گره‌های تعیین‌کننده در یک شبکه شناسایی می‌شوند تا با استفاده از آن‌ها، به شناسایی کانال‌های کلیدی یک شبکه به منظور تبلیغات اثرگذارتر پرداخته شود [۲۱].

با این حال، تمامی تحقیقات انجام‌شده در زمینه بازاریابی ویروسی، با شکل رهگیری نتایج و رسیدن به سرشاخه‌ها نیست. در واقع، برخی محققان اعتقاد دارند که رفتار تصادفی برخی از کاربران و یا احتمال ایجاد یک رفتار تصادفی در تمامی کاربران، می‌تواند باعث تغییر الگوی انتشار یک اطلاعات گردد و در نتیجه، در واقعیت، امکان رهگیری یک اطلاع و شناسایی منبع انتشار آن را ناممکن می‌سازد. در این حالت، محققین از روش‌های غیرقطعی و احتمالی استفاده می‌کنند که در آن، عبور اطلاعات از هر گره و انتشار آن، به صورت تصادفی است. برای تعیین احتمال عبور، توابع مختلفی ارائه شده است که شامل انواع معیارهای مرکزیت، همسایگی و روش‌های خوشه‌بندی می‌گردد [۲۲].

چالش مشترک در تمامی تحقیقات بازاریابی ویروسی، حجم زیاد اطلاعات و عدم امکان پردازش این اطلاعات است. به همین جهت، معیارهای سراسری مرکزیت، معمولاً توانایی پردازش شبکه‌های بزرگ را از دست می‌دهند. با این حال، معیارهای محلی نیز توانایی ارائه نتایج دقیق به اندازه معیارهای سراسری را ندارند. بنابراین، یک روش بینابینی که مورد استفاده برخی تحقیقات قرار می‌گیرد؛ توسعه معیارهای مرکزیت نیمه محلی است. در استفاده از این معیارها، بخش بزرگ‌تری از شبکه به نسبت به معیارهای محلی مورد تحلیل قرار می‌گیرد و سرعت پردازش اطلاعات در آنها نیز بیشتر از معیارهای سراسری است [۲۳].

با وجود معیارهای مختلف توسعه داده شده، مسأله گستردگی شبکه‌های اجتماعی و حجم بالای اطلاعات، همچنان محل ابهام است. در همین راستا، روش‌های خوشه‌بندی دینامیک داده کاوی، در این تحقیقات مورد استفاده قرار گرفته است. در این روش‌ها، بر مبنای وضعیت موجود سیستم، کاربران در خوشه‌های مختلف طبقه‌بندی می‌شوند و تعداد کاربرانی که باید مورد مطالعه دقیق قرار گیرند؛ کاهش می‌یابد. بدین

ترتیب، می‌توان محدوده مورد بررسی را کاهش داده و هم‌زمان، سرخوشه‌های با اهمیت را با دقت بالایی شناسایی کرد [۲۳].

در تحقیق حاضر، با توجه به اهمیت استفاده از بازاریابی ویروسی برای تبلیغات و رقابت در بازارهای الکترونیکی، از یک روش ترکیبی خوشه‌بندی و مرکزیت نیمه محلی استفاده شده است که در آن، برای خوشه‌بندی دینامیک کاربران یک شبکه، از معیار مرکزیت نیمه محلی استفاده می‌شود. خوشه‌بندی دینامیک، با استفاده از روش زنجیره مارکوف انجام می‌پذیرد و برای تعیین مرکزیت نیمه محلی نیز از روابط موجود بین گره‌ها در شبکه‌های اجتماعی شامل لایک و کامنت، استفاده خواهد شد.

۳- روش تحقیق

به‌طورکلی، اگر هر کاربر در شبکه مجازی به صورت یک گره و هرگونه ارتباط بین کاربران به‌وسیله یک یال مشخص گردد؛ شبکه مجازی به صورت یک گراف بسیار بزرگ قابل نمایش است. ارتباط در شبکه‌های مجازی، غالباً از طریق ابزارهایی همچون لایک و کامنت ایجاد می‌گردد و لذا هرچه تعداد لایک و کامنت بین دو کاربر بیشتر باشد؛ ارتباط قوی‌تری بین دو کاربر وجود دارد. بنابراین، نکته قابل توجه این است که این گراف، یک گراف وزن‌دار جهت‌دار نیز است.

- **خوشه‌بندی گراف**

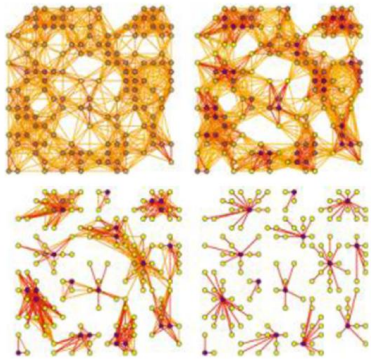
در مقدمه تحقیق گفته شد که در این پژوهش، به منظور شناسایی کاربران کلیدی در شبکه‌های اجتماعی، از روش خوشه‌بندی گراف استفاده می‌شود. به صورت کلی، منظور از خوشه‌بندی گراف، قراردادن گره‌های مشابه از یک گراف در یک خوشه است. گره‌های مشابه، معمولاً گره‌های نزدیک به یکدیگر است و لذا فاصله هندسی گره‌ها، می‌تواند معیار خوبی برای خوشه‌بندی باشد. با استفاده از معیارهای مختلف فاصله هندسی همچون گریز از مرکز، شعاع و قطر به خوشه‌بندی گره‌ها پرداخته است که به این منظور، از الگوریتم نزدیک‌ترین همسایگی استفاده کرده است. نکته مهم در استفاده از فواصل هندسی برای خوشه‌بندی، این است که این فواصل باید به خوبی قابل تعریف بوده و در واقع دوری و نزدیکی با استفاده از یک معیار مشخص، قابل تعریف باشد. این در حالی است که فاصله مکانی با استفاده از شبکه‌های مجازی و اینترنت، تفاوتی بین کاربران مختلف ایجاد نکرده و لذا نمی‌توان فواصل بین گره‌ها را با معیار خوبی محاسبه نمود [۲].

در چنین شرایطی، می‌توان به خوشه‌بندی با استفاده از روش قدم‌زدن تصادفی و به‌خصوص با استفاده از زنجیره مارکوف پرداخت. در روش قدم‌زدن تصادفی، شباهت گره‌ها با یکدیگر نه براساس فاصله فیزیکی، که براساس رفتارهای ارتباطی آن‌ها سنجیده می‌شود. برای مثال، در صورتی که دو گره به یکدیگر متصل باشند و این ارتباط وزن بالایی نیز داشته باشد؛ احتمال قرارگرفتن آن‌ها در یک خوشه افزایش می‌یابد؛ در صورتی‌که ارتباط مستقیم وجود نداشته باشد، در صورتی‌که دو گره به یک گره مشترک متصل باشند؛ باز هم شانس قرارگرفتن در یک خوشه را دارند و این ارتباط تا سطوح بعدی نیز پیش رفته و در واقع تمامی ارتباطات چندسطحی ممکن را در نظر می‌گیرد. به این ترتیب، می‌توان معیار کلی شباهت دو گره در روش قدم‌زدن تصادفی را مشابه رابطه (۱) دانست.

$$s(u,v)=\frac{C}{|I(u)||I(v)|}\sum_{x\epsilon u}\sum_{y\epsilon v}s(x.y) \qquad (۱)$$

در این رابطه، شباهت دو گره x و y براساس تعداد همسایه‌های مشابه در مجموعه همسایگی‌های مربوطه که به ترتیب u و v را شامل می‌شوند؛ محاسبه می‌گردد. به این ترتیب، این روش، با محاسبه شباهت‌های گره‌های مختلف گراف به صورت دو به دو، به پیداکردن گره‌های مشابه و خوشه‌بندی این گره‌ها می‌پردازد.

زنجیره مارکوف، یکی از زیرمجموعه‌های روش قدم‌زدن تصادفی است که با استفاده از منطق شبیه‌سازی به خوشه‌بندی گره‌ها در یک گراف می‌پردازد. منطق کار این روش، این است که از یک گره شروع کرده و به صورت احتمالی به یکی از گره‌های مشابه حرکت می‌کند؛ این کار را برای تعداد مراحل مشخصی انجام داده و در نهایت، گره‌هایی که از آن‌ها عبور شده است؛ در یک خوشه قرار می‌گیرند؛ همچنین تعداد عبورهای صورت‌گرفته از هر گره نیز تعیین می‌گردد. دو نکته کلیدی در استفاده از این روش وجود دارد؛ اول اینکه احتمال حرکت از یک گره به گره‌های مجاور، یکسان نبوده و برای هریک از گره‌های مجاور وزنی در نظرگرفته می‌شود. دوم ممکن است یک گره در چندین خوشه قرار گیرد که در این حالت، باید گره مربوطه به خوشه/ای تخصیص یابد که بیشترین عبور از آن در آن خوشه اتفاق افتاده است. به عبارت دیگر، خوشه‌های ایجادشده توسط زنجیره مارکوف، در ابتدا همپوشانی زیادی با یکدیگر دارند؛ لیکن با تعداد بیشتر اجرای الگوریتم و تخصیص بهینه گره‌های مشابه، این همپوشانی کاهش یافته و در نهایت به صفر می‌رسد. شکل ۱ این موضوع را به‌خوبی نشان می‌دهد.



شکل ۱- سیر تغییرات خوشه‌بندی به روش مارکوف و از بین رفتن همپوشانی خوشه‌ها

به این ترتیب، مشکل گره‌های مشابه در خوشه‌های مختلف، رفع می‌گردد. اما در مورد محاسبه احتمال، روش‌های گوناگونی وجود دارد. در تحقیقات مختلف، معیارهای متفاوتی برای سنجش احتمال حرکت از یک گره به گره دیگر استفاده شده است. متداول‌ترین معیار، وزن یال متصل‌کننده دو گره است که هر چه میزان ارتباط دو گره بیشتر باشد؛

احتمال حرکت به آن گره بیشتر می‌شود. لیکن در تحقیق اخیر، از یک معیار جدید برای خوشه‌بندی به این روش استفاده می‌گردد. این معیار، میزان مرکزیت گره‌های مجاور است؛ در واقع، هرچه میزان مرکزیت یک گره افزایش یابد، احتمال حرکت به آن گره نیز افزایش پیدا می‌کند. محقق، اعتقاد دارد که در شبکه‌های اجتماعی، گاهی نمی‌توان ارتباط مستقیمی بین دو گره دریافت؛ لیکن می‌توان انتظار داشت که گره‌های با مرکزیت بالاتر، مخاطب بیشتری داشته باشند؛ به عبارت دیگر، ممکن است کاربری از طریق ابزارهای لایک، کامنت یا سایر ارتباطات مجازی، ارتباط مستقیمی با کاربر دیگر نداشته باشد؛ اما در صورت بالابودن مرکزیت آن کاربر، احتمال مشاهده صفحات مربوط به وی و تأثیرگرفتن از او، افزایش می‌یابد. به همین دلیل، در این پژوهش از معیار مرکزیت نیمه محلی برای خوشه‌بندی استفاده می‌شود که در ادامه این قسمت به توضیح آن پرداخته می‌شود.

- **مرکزیت نیمه محلی**

مرکزیت به صورت کلی، یکی از معیارهای سنجش میزان ارتباطات یک گره، با گره‌های دیگر است و هرچه مقدار آن بیشتر باشد؛ به معنای تأثیرگذاری بیشتر آن گره در گراف است. برای اندازه‌گیری مرکزیت، دو حالت مرکزیت سراسری و مرکزیت محلی وجود دارد. در محاسبه مرکزیت سراسری، تمامی گره‌های گراف در نظر گرفته شده و برای محاسبه مرکزیت محلی، تنها گره‌های مجاور در سطح اول یا چند سطح ابتدایی در نظر گرفته می‌شوند. هر دو روش، مزایا و معایب خاص خود را دارند؛ برای مثال، دقت مرکزیت سراسری از محلی بیشتر است؛ لیکن محاسبات مربوط به آن زیادتر و در مورد شبکه‌های بزرگ، عملاً غیرممکن می‌گردد. در مقابل، مرکزیت محلی، تخمینی از مرکزیت سراسری برای هر گره بوده و محاسبات آن با حجم کمتر و سرعت بالاتر صورت می‌پذیرد. در تحقیق حاضر، به منظور استفاده از مزایای هر دو روش، از یک روش ترکیبی سنجش مرکزیت با نام مرکزیت نیمه محلی استفاده شده است که در ادامه به توضیح آن پرداخته می‌شود.

رابطه (۲) نشان‌دهنده مبنای محاسبه مرکزیت محلی هر گره است.

$$C_i^w = \frac{d_i}{n-1} \qquad (۲)$$

در این رابطه، d_i نشان‌دهنده وزن مجموع یال‌های وارد شونده به گره i است که به نسبت وزن مجموع تمامی یال‌های گراف، نرمال می‌گردد. مقدار مرکزیت ابتدایی هرگره، توسط رابطه (۳) به مرکزیت نهایی هر گره تبدیل می‌گردد.

$$M_j^w = \sum_{j\epsilon u} C_j^w \qquad (۳)$$

در رابطه مذکور، مجموعه u مجموعه نزدیک‌ترین همسایگی گره i است و در واقع میزان مرکزیت نهایی هر گره، برابر مرکزیت همان گره به علاوه مرکزیت گره‌های مجاور است. با تعدیل میزان مرکزیت به نسبت ویژگی‌های فردی، معیار مرکزیت تعدیل شده هر گره مطابق رابطه (۴) محاسبه می‌شود.

$$Q^w = \sum_{j\epsilon u} (M^w(j) * (S + A + G)) \qquad (۴)$$

در این رابطه، S مقدار نرمال‌شده جنسیت است؛ در واقع اگر کاربر متناظر یک گره جنسیت مرد داشته باشد، مقدار نرمال‌شده برابر تعداد کل مردان به تعداد کل جمعیت شبکه است. متغیر A به صورت پیوسته و نشان‌دهنده سن نرمال‌شده هرگره و G نشان‌دهنده امتیاز مدرک تحصیلی فرد است. در نهایت، مقدار مرکزیت نیمه محلی هریک از گره‌های حاضر در شبکه مطابق رابطه (۵) تعریف می‌شود.

$$DDSC(v) = \sum_{u\epsilon v} w_{vu}\, Q^w(u) \qquad (۵)$$

در این رابطه، مقدار مرکزیت نیمه محلی هرگره، برابر وزن یال‌های ورودی هرگره در میزان مرکزیت تعدیل شده همان گره است.

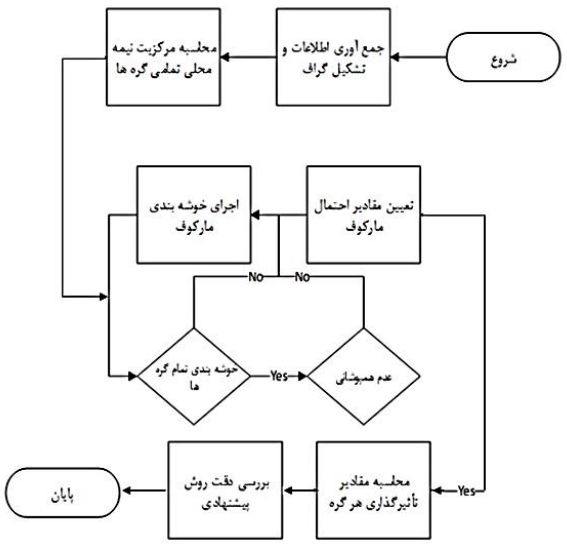
با توجه به مراحل فوق، می‌توان مقدار مرکزیت نیمه محلی مربوط به هرگره را محاسبه نمود و این مسأله، مقادیر مورد نیاز برای استفاده از روش خوشه‌بندی پیشنهادی تحقیق را فراهم می‌کند.

- **خوشه‌بندی**

پیشتر گفته شد که برای خوشه‌بندی در این تحقیق، از روش خوشه‌بندی مارکوف استفاده می‌شود که معیارهای مربوطه برای خوشه‌بندی، مقادیر مرکزیت نیمه محلی به‌دست آمده تعریف می‌شود. تنها مسأله باقیمانده، محاسبه احتمال حرکت از هر گره به گره دیگر است که این مسأله براساس رابطه (۶) انجام می‌پذیرد.

$$p(u,v) = \frac{DDSC_v}{\sum_v DDSC_v} \qquad (۶)$$

براساس این رابطه، از میان گره‌های مجاور، احتمال حرکت به گره‌ای که بیشترین مرکزیت را دارد بیشتر از سایر گره‌ها است و مجموع احتمالات حرکت از هر گره به گره‌های مجاور نیز، برابر ۱ است. با توجه به مطالب توضیح داده شده در این قسمت، به معرفی مدل مفهومی تحقیق مطابق شکل ۲ پرداخته می‌شود.



شکل ۲- مشخصات کلی شاخص‌های تحقیق

براساس شکل ۲، پس از محاسبه مرکزیت نیمه محلی تمامی گره‌ها، خوشه‌بندی به روش مارکوف انجام می‌گردد. اما همچنان یک سؤال دیگر باقی مانده است و آن چگونگی محاسبه میزان اثرگذاری هر گره با استفاده از نتایج خوشه‌بندی است. همان‌طور که گفته شد؛ در خوشه‌بندی مارکوف، تعداد گذر ثبت شده از هر گره ذخیره می‌شود. براساس تعداد بار عبور از هر گره در هر خوشه و تعداد گره‌های هر خوشه، می‌توان میزان اثرگذاری را مطابق رابطه ۷ محاسبه کرد.

$$E_u = NoP_u * \frac{NoK_c}{\sum_c NoK_c} \quad (۷)$$

در رابطه (۷)، NoP_u تعداد بارهای عبور از گره u است. همچنین، نسبت حاضر در قسمت دوم نشان‌دهنده نسبت تعداد گره‌های حاضر در خوشه مربوط به u به نسبت تعداد کل گره‌های موجود است. به این ترتیب، هرچه تعداد گذر از گره‌ای بالاتر بوده و تعداد گره‌های حاضر در خوشه مربوطه نیز بیشتر باشد؛ میزان اثرگذاری گره در شبکه بیشتر می‌شود و از نظر منطق بازاریابی ویروسی، این گره یک گره مهم تلقی می‌گردد.

- **سنجش کارایی**

به منظور سنجش کارایی روش پیشنهادی، همانند هر تحقیق دیگر باید به مقایسه نتایج آن با یک مقدار واقعی یا مرجع پرداخت. در تحقیقات مختلف، معمولاً از مدل SIR برای تخمین مقادیر واقعی اهمیت کاربران و رتبه‌بندی آن‌ها استفاده می‌شود. چگونگی محاسبه این معیار مطابق رابطه (۸) است.

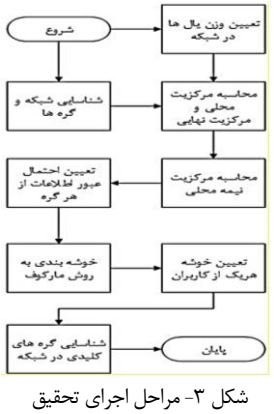
$$\gamma_{ij} = (\frac{w_{ij}}{w_m})^{\beta} \quad (۸)$$

عملکرد این مدل به این صورت است که هر بار یکی از گره‌ها در شبکه انتخاب شده و الگوریتم در مورد آن اعمال می‌گردد. الگوریتم، با استفاده از شبیه‌سازی به بررسی توانایی انتشار هریک از گره‌ها می‌پردازد.

رابطه (۸) نشان‌دهنده احتمال سرایت آلودگی از گره i به گره j است که شامل نسبت وزن یال به وزن گره به توان مقدار پیش فرض است.

به منظور سنجش کارایی روش پیشنهادی، به مقایسه همبستگی رتبه‌های نهایی به‌دست آمده با رتبه‌بندی ایجادشده توسط SIR با استفاده از آماره معناداری تاوکندل پرداخته خواهد شد.

به این ترتیب، می‌توان مراحل اجرایی تحقیق را مطابق شکل ۳ دانست.



شکل ۳- مراحل اجرای تحقیق

بر مبنای شکل نمایش داده شده، ابتدا یک نمونه آماری از شبکه اجتماعی اینستاگرام جمع‌آوری شده و شبکه اجتماعی تشکیل می‌شود. در این شبکه نمونه، کاربران، گره‌های شبکه و ارتباطات لایک و کامنت، یال‌های شبکه به حساب می‌آیند. برای محاسبه مرکزیت نیمه محلی، ابتدا باید معیارهای مرکزیت محلی، مرکزیت نهایی و مرکزیت نهایی تعدیل شده محاسبه شوند. برای محاسبه تمامی این معیارها، از تعداد ارتباطات گره‌ها با یکدیگر و تعداد گره‌های تماس با یک گره مشخص، استفاده می‌شود.

با تعیین مقدار مرکزیت نیمه محلی برای هریک از گره‌های شبکه نمونه، احتمال انتشار اطلاعات از این گره به هریک از گره‌های مجاور محاسبه می‌شود.

در ادامه، با شبیه‌سازی یک اطلاعات، اطلاعات جدیدی به شبکه وارد می‌شود و از طریق گره‌ها انتشار می‌یابد. با هر بار انتشار، گره‌ها توسط روش مارکوف خوشه‌بندی می‌شوند. پس از انجام تکرارهای متوالی از ایجاد اطلاعات در گره‌های مختلف، بر مبنای تعداد بارهایی که یک گره اطلاعات ایجاد کرده و میزان پوشش ایجاد شده اطلاعات عبوری از این گره، یک امتیاز به این گره اختصاص می‌یابد.

در نهایت، با استفاده از امتیازات به‌دست آمده، کاربران با اهمیت و گره‌های حیاتی شبکه شناسایی می‌شوند. همچنین، به منظور بررسی کارایی این روش ترکیبی، از مقایسه میزان همبستگی نتایج به‌دست آمده از این روش با سایر معیارهای مرکزیت محلی و نیمه محلی، با معیار SIR استفاده خواهد شد.

۴- نتایج

به منظور بررسی روش پیشنهادی تحقیق، یک نمونه آماری ۱۰۰ نفره در شبکه اجتماعی اینستاگرام انتخاب شده و رفتارهای این شبکه براساس دو معیار لایک و کامنت به صورت جداگانه مورد بررسی قرارگرفته است. خصوصیات کلی نمونه انتخاب شده همانند جدول ۱ است.

جدول ۱- مشخصات کلی شاخص‌های تحقیق

شاخص	مجموع	میانگین	انحراف‌معیار
تعداد کامنت	۱۶۶۹۴۹	۲۵/۴۹	۳۲/۵۴
تعداد لایک	۲۵۴۹۱۶	۱۶/۶۹	۲۸/۹۳
سن	-	۲۲	۶/۴۷
تحصیلات	-	۱/۴۵	۰/۳۴
جنسیت	-	۰/۵۷	۰/۲۸

شاخص جنسیت، با اختصاص عدد ۰ به خانم و عدد ۱ به آقا اندازه‌گیری شده است. شاخص تحصیلات نیز با اختصاص عدد ۰ به بی‌سواد، عدد ۱ به تحصیلات قبل از دانشگاه و عدد ۲ به تحصیلات دانشگاهی اندازه گرفته شده است. با توجه به جدول فوق، مشخص می‌شود که بدنه اصلی شبکه نمونه را جوانان ۲۲ ساله، دارای تحصیلات دانشگاهی و مرد ایجاد کرده است. همچنین، تعداد لایک بیش از ۵۰٪ از تعداد کامنت بیشتر بوده و انحراف‌معیار آن نیز زیادتر است.

به منظور اندازه‌گیری اهمیت کاربران در شبکه، علاوه بر روش پیشنهادی، از روش‌های دیگری به‌عنوان رقیب برای روش پیشنهادی استفاده شده است. به همین منظور، خصوصیات آماری مقادیر محاسبه شده برای اهمیت کاربران در جدول ۲ مقایسه شده‌اند. لازم به ذکر است که مقادیر نشان داده شده در این جدول، به صورت میانگین مقادیر مربوط به معیارهای لایک و کامنت است که در این جدول خلاصه شده و در قالب یک جدول نمایش داده می‌شود.

جدول ۲- مقایسه آماره‌های توصیفی سنجه‌های مختلف اهمیت کاربران در شبکه مجازی نمونه

سنجه	میانگین	انحراف‌معیار	چولگی	کشیدگی
مرکزیت درجه	۱۶/۶۴	۲۸/۹۸	-۰/۲۴	۲/۵۸
مرکزیت نزدیکی	۰/۱۵	۰/۱۱	-۰/۵۳	۴/۵۷
مرکزیت محلی	۰/۵۱	۰/۲	۰/۲۸	۲/۹۹
مرکزیت نیمه محلی	۴۰/۸۸	۳۲/۷۱	۲/۱۹	۴/۹۶
مرکزیت خوشه‌بندی (روش پیشنهادی)	۰/۹۳	۰/۶۵	۰/۳۲	۳/۱۸

به صورت خلاصه، می‌توان اطلاعات زیر را از جدول ۲ استخراج نمود.

۱. مرکزیت درجه: این معیار، نشان‌دهنده تعداد گره‌های متصل به هر گره است و لذا آماره‌های نشان داده شده برای آن، شبیه به آماره‌های توصیفی داده‌های اصلی در جدول ۱ است. چولگی این سنجه کمتر از ۰.۵ بوده و قابل اغماض است و کشیدگی آن نیز در حدل نرمال است.

۲. مرکزیت نزدیکی: این معیار، عکس فاصله گره تا گره‌های مجاور خود را نشان می‌دهد و برای همین کمتر از ۱ تخمین زده شده است. چولگی منفی شناسایی‌شده در آن معنادار بوده و کشیدگی داده‌ها نیز از حد نرمال بالاتر است. بنابراین، سنجش اهمیت با این معیار باعث تخمین مقادیر بسیار نزدیک در مورد اکثر کاربران و مقادیر بسیار کم در مورد تعدادی از کاربران می‌گردد و نمی‌تواند معیار مناسبی برای تخمین اهمیت کاربران باشد.

۳. مرکزیت محلی: این معیار، همانند مرکزیت درجه بوده و در محاسبه تعداد یال‌های هر گره، تعداد یال‌های گره‌های مجاور به نسبت کلیه یال‌ها را در نظر می‌گیرد. همانند معیار مرکزیت درجه نیز دارای چولگی و کشیدگی در حد نرمال است.

۴. مرکزیت نیمه محلی: این معیار که چگونگی محاسبه آن توضیح داده شده است؛ دارای چولگی مثبت قابل توجه و کشیدگی بیش از حد نرمال است که این امر نشان‌دهنده کارایی ضعیف این معیار در سنجش اهمیت و رتبه‌بندی کاربران شبکه است؛ چرا که تعداد زیادی از کاربران مقادیر بسیار نزدیک و تعداد کمی مقادیر بسیار بزرگ به خود می‌گیرند.

۵. مرکزیت خوشه‌بندی: این معیار، که مطابق روش پیشنهادی تحقیق و رابطه ۸ به‌دست می‌آید؛ دارای چولگی و کشیدگی در حد نرمال بوده و انحراف‌معیار آن نیز مقدار معقولی به خود گرفته است و در نتیجه، می‌تواند جهت رتبه‌بندی نتایج مناسبی تولید نماید.

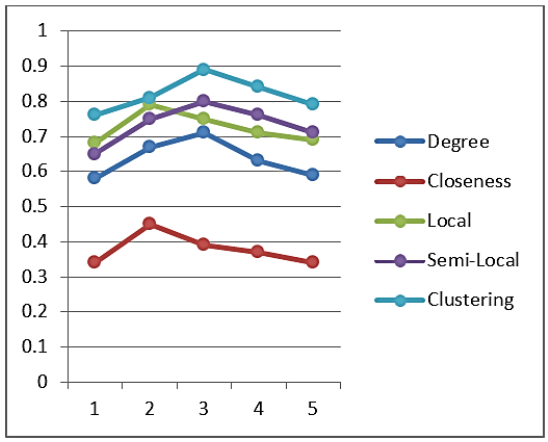
حال، به منظور سنجش میزان ارتباط سنجه پیشنهادی تحقیق با استفاده از خوشه‌بندی، با سایر معیارهای به‌کار گرفته شده در جدول ۲، از آزمون همبستگی اسپیرمن استفاده می‌شود که نتایج آن در جدول ۳ نشان داده شده است.

جدول ۳- بررسی معناداری رابطه میان سنجه پیشنهادی با سایر سنجه‌ها با استفاده از آزمون اسپیرمن

سنجه	آماره اسپیرمن	معناداری	سطح معناداری
مرکزیت درجه	۱/۹۵	۰/۰۲	٪۹۵
مرکزیت نزدیکی	۰/۸۶	۰/۱۹	-
مرکزیت محلی	۲/۱۷	۰/۰۱	٪۹۹
مرکزیت نیمه محلی	۱/۵۷	۰/۰۵	٪۹۰

براساس جدول ۳، همبستگی روش پیشنهادی با مرکزیت نزدیکی معنادار نیست. مرکزیت محلی بیشترین همبستگی و مرکزیت درجه و نیمه محلی رتبه‌های بعدی همبستگی را دارند.

در نهایت، برای سنجش کارایی معیار پیشنهادی، از مقایسه همبستگی آن با شاخص SIR و با همبستگی سایر سنجه‌ها با این شاخص، استفاده می‌گردد. نتایج این مهم، در شکل ۳ آورده شده است.



شکل ۴- مقادیر همبستگی سنجه‌های مختلف اهمیت کاربران با شاخص SIR به ازای بتای متفاوت

۵ آزمایش نشان داده شده در شکل ۳ به ازای مقادیر مختلف بتا (توان شاخص SIR) است که این توان نشان‌دهنده میزان احتمال حرکت و در نتیجه روانی حرکت است. با توجه به شکل، مشخص است که همبستگی شاخص پیشنهادی (Clustering) از مابقی شاخص‌ها بالاتر بوده است. همچنین، مقدار حداکثر همبستگی در آزمایش سوم به‌دست می‌آید که نشان‌دهنده مقدار ۱ برای بتا است و مقادیر کمتر یا بیشتر از آن، باعث کاهش همبستگی می‌گردد. مقادیر همبستگی نشان داده شده، براساس رتبه‌بندی انجام‌شده توسط روش‌ها بوده و توسط آماره تاوکندل به‌دست آمده است.

۵- جمع‌بندی

با توجه به ضرورت شناسایی کاربران با اهمیت در شبکه‌های اجتماعی به منظور تبلیغات مؤثر در چارچوب بازاریابی ویروسی، تحقیق حاضر به ارائه یک مدل ترکیبی خوشه‌بندی جهت تخمین اهمیت کاربران پرداخته است. روش پیشنهادی، استفاده از زنجیره مارکوف جهت خوشه‌بندی کاربران است که البته برخلاف روش مارکوف معمولی که از وزن یال‌ها استفاده می‌کند؛ از میزان مرکزیت نیمه محلی به منظور محاسبه احتمال حرکت استفاده می‌کند. همچنین، اهمیت کاربران در شبکه به اهمیت کاربر در خوشه و اهمیت خوشه در گراف بستگی دارد. براساس نتایج به‌دست آمده، توزیع مقادیر اهمیت سنجش شده به صورت نرمال بوده و این مسأله، مناسب‌بودن مقادیر جهت رتبه‌بندی را تأیید می‌کند. علاوه بر این، براساس آزمون همبستگی صورت گرفته با معیار تاوکندل، معلوم شد که بیشترین همبستگی مربوط به روش پیشنهادی با روش SIR است و لذا کارایی روش ترکیبی در تخمین مقادیر SIR نسبت به سایر روش‌ها بیشتر است. همچنین، از میان سایر معیارهای مرکزیت، همبستگی روش پیشنهادی با روش مرکزیت محلی بیشتر از سایر روش‌ها بوده و از لحاظ آماری معنادار بوده است. پیش از این و در [۳] اثبات شده بود که همبستگی شاخص نیمه محلی استفاده شده در این تحقیق، با معیار SIR بیشتر از سایر معیارهای نیمه محلی محلی بوده است که البته این موضوع در این تحقیق نیز مورد تأیید قرار گرفته است.

علاوه بر این، [۲] نیز با تکیه بر مزایای استفاده از یک روش دینامیک برای خوشه‌بندی گره‌های یک شبکه اجتماعی، به این نتیجه رسیده است که کارایی روش مارکوف در میان روش‌های دسته‌بندی و روش‌های استاتیک، بالاتر است که این یافته نیز منطبق بر نتایج این تحقیق است.

بنابراین، می‌توان گفت که تحقیق حاضر با ترکیب دو روش مورد استفاده قرار گرفته به‌عنوان روش بهینه در سایر تحقیقات، توانسته است روشی برای خوشه‌بندی دینامیک کاربران در یک شبکه اجتماعی پیدا کند که می‌تواند مورد استفاده‌های زیادی قرار گیرد. با تحلیل شبکه‌های ارتباطی مشتریان، شرکت‌ها می‌توانند اطلاعات مورد نظر خود و پخش آن در شبکه را کنترل کنند و آن را به حد مطلوب مدنظر خود برسانند. افزایش کارایی بازاریابی جهت افزایش فروش محصولات نیز از دیگر کاربردهای محتمل روش پیشنهادی این تحقیق است. با توجه به دینامیک‌بودن کارکرد این روش، می‌توان آن را برای شبکه‌های بزرگ‌تر نیز به‌کار برد و همچنان، با دقت بالایی مهم‌ترین گره‌های یک شبکه را شناسایی کرده و بیشینه پوشش مدنظر را ایجاد کرد.

۶- مراجع

۱- بیژن، علمی؛ مجتبی، رمضانی، بررسی جامعه‌شناختی تاکتیک‌های مؤثر بر بازاریابی ویروسی و نقش آن بر رفتار خرید، مجله مطالعات جامعه‌شناسی، شماره ۱۰، صفحات ۱۲۵ تا ۱۴۴، ۱۳۹۰.

۲- محمدعلی، محمودیار، یک مدل مبتنی بر گراف برای ارزیابی اعتماد در جوامع برخط، پایان‌نامه جهت اخذ مدرک کارشناسی‌ارشد، دانشکده مهندسی صنایع، دانشگاه صنعتی خواجه نصیرالدین طوسی، ۱۳۹۵.

۳- احد، جعفرزاده، ارائه مدل بازاریابی ویروسی در شبکه‌های اجتماعی با استفاده از تئوری گراف، پایان‌نامه جهت اخذ مدرک کارشناسی‌ارشد، دانشکده مهندسی صنایع، دانشگاه صنعتی خواجه نصیرالدین طوسی، ۱۳۹۵.

۴- سوسن، روح‌روان؛ منیره، حسینی، بررسی فرصت‌های بازاریابی ویروسی موبایل در ایران، هفتمین همایش ملی و اولین همایش بین‌المللی تجارت و اقتصاد الکترونیکی، تهران، انجمن علمی تجارت الکترونیک ایران، ۱۳۹۲.

۵- نسترن، حاجی حیدری؛ امیر، خانلری؛ حمیدرضا، ریحانی، شناسایی و تعیین اهمیت عوامل مؤثر بر بازاریابی ویروسی در حوزه نرم‌افزارهای کاربردی موبایل، مدیریت فناوری اطلاعات، ۹(۲)، ۲۳۷ تا ۲۵۲، ۱۳۹۶.

6- Kaple. M, Kulkarni. K, Potika. K, Viral Marketing for Smart Cities: Influencers in Social Network Communities, IEEE Third International Conference on Big Data Computing Service and Applications, **DOI:** 10.1109/BigDataService. 2017.46, 2017.

7- Hung T. Nguyen, My T. Thai, and Thang N. Dinh, A Billion-Scale Approximation Algorithm for

8- Maximizing Benefit in Viral Marketing, IEEE/ACM TRANSACTIONS ON NETWORKING, No. 181, Vol. 23, 2017.

9- Li. Y. M, Lai. C. Y, Chen. C. W, Discovering Influencers for Marketing in the Blogosphere, Information Sciences, Vol. 25, Issue. 4, pp. 5143-5157, 2011.

10- Kaplan, A. M, Haenlein. M, Two Hearts in Three-Quarter Time: How to Waltz the Social Media/Viral Marketing Dance, Business Horizons, 54(3), 253-263, 2011.

11- Zhuo-ming. R, Jian-guo. L, Feng. S, Zhao-long. H and Qiang. G, Analysis of the Spreading Influence of the Nodes with Minimum K-shell Value in Complex Networks. Acta Phys. Sin, 62 (10), Article 108902, 2013.

12- Zeng. A and Zhang, C. J, Ranking Spreaders by Decomposing Complex Networks, Physics Letters A, 377 (14), pp. 1031–1035, 2013.

13- Berger. J, Iyengar. R, How Interest Shapes Word-of- Mouth Over Different Channels, The Wharton School of the University of Pennsylvania, 2012.

14- Berger. J and Milkman. K. L, What Makes Online Content Viral? Journal of Marketing Research, 49(2), 192–205, 2012.

15- Long. C, Wong. R. C-W, Viral marketing for dedicated customers. Information Systems, 46, 1-23, 2014.

16- Oliveira. M, Gama. J, An Overview of social Network Analysis, Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, Vol.2, No.2, pp.99-115, 2012.

17- Han. J, Pei. J, Kamber. M, Data Mining: Concepts and Techniques, Third Edition , Elsivier, MA, USA, 2011.

18- Bauer. f, Lizier. J. T, Identifying Influemtial Speareders and Efficiently Estimating Infection Numbers in Epidemic Models: A Walk Counting Approach, EPL Europhysical Letters. No.99, Vol.6, pp.68007, 2012.

19- Rad. A, Benyoucef. M, Towards Detecting Influential Users in social Networks, E- Technologies: Transformation in a Connected World, 5th International Conference, pp. 227–240, 2011.

20- Rodrigues, H. S., & Fonseca, M. J. (2016). Can information be spread as a virus? Viral marketing as epidemiological model. Mathematical Methods in the Applied Sciences, 39(16), 4780-4786.

21- Ahmed, H. M. S. (2018). A Proposal Model for Measuring the Impact of Viral Marketing Through Social Networks on Purchasing Decision: An Empirical Study. International Journal of Customer Relationship Marketing and Management (IJCRMM), 9(3), 13-33.

22- Wang, Y., Vasilakos, A. V., Ma, J., & Xiong, N. (2015). On studying the impact of uncertainty on behavior diffusion in social networks. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 45(2), 185-197.

23- Kaple, M., Kulkarni, K., & Potika, K. (2017, April). Viral marketing for smart cities: Influencers in social network communities. In 2017 IEEE Third International Conference on Big Data Computing Service and Applications (BigDataService) (pp. 106-111). IEEE.

24- Helal, N. A., Ismail, R. M., Badr, N. L., & Mostafa, M. G. (2016). A novel social network mining approach for customer segmentation and viral marketing. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 6(5), 177-189.