مجله علمی پژوهشی «پدافند الکترونیکی و سایبری»
سال .....، شماره ....، فصل ۱۴۰۰؛ ص

مدل محاسباتی جهت ارزیابی عملکرد عامل عملیات نفوذ در شبکه‌های اجتماعی برخط

غلامرضا بازدار
۱. دانشجوی دکتری، دانشگاه امام حسین (ع)، تهران، ایران

محمد عبداللهی ازگمی*
استاد، دانشگاه علم و صنعت، تهران، ایران

(دریافت: ۱۴۰۲/۱۰/۱۱، بازنگری: ۱۴۰۳/۰۲/۰۸، پذیرش: ۱۴۰۲/۰۲/۱۹، انتشار: ۱۴۰۳/۰۳/۱۳)

1- DOR: https://dorl.net/dor/

چکیده

گسترش شتابان استفاده از شبکه‌های اجتماعی برخط در میان جامعه، زمینه مستعدی برای اجرای عملیات‌های نفوذ شناختی و اجتماعی را فراهم آورده است. طرح‌ریزی و اجرای بهینه عملیات نفوذ، وابسته به داشتن یک چارچوب مناسب جهت ارزیابی این عملیات است. ارزیابی عامل‌ها و بازیگران مؤثر در عملیات نفوذ از ملزومات اصلی ارزیابی عملیات نفوذ است. با توجه به پویایی شبکه‌های اجتماعی برخط و تولید روزافزون داده‌های انبوه در آن، استفاده از رویکرد محاسباتی جهت ارزیابی عملیات نفوذ ضروری است. لذا هدف این تحقیق، یافتن مدلی محاسباتی جهت ارزیابی عامل‌های عملیات نفوذ در شبکه‌های اجتماعی برخط است. به‌طورکلی روش‌های ارزیابی نفوذ عامل را می‌توان به سه دسته کلی ارزیابی کیفی، ارزیابی کمی و ارزیابی محاسباتی تقسیم کرد. روش‌های ارزیابی محاسباتی را می‌توان به دو دسته‌ی روش‌های مبتنی بر یادگیری ماشین و روش‌های مبتنی بر ویژگی‌های دستی[1] تقسیم‌بندی کرد. روش‌های دسته‌ی اول دارای دقت بالاتری هستند اما نیاز به حجم زیادی داده‌ی آموزش دارند. این در حالی است که در مسائلی همچون مسئله رتبه‌بندی نفوذ عامل‌ها، امکان آماده‌سازی داده‌های برچسب‌دار وجود ندارد. یکی دیگر از معایب غالب روش‌های مبتنی بر یادگیری ماشین، عدم امکان قابلیت تفسیر نتایج است. همچنین با بهره‌گیری از سازه‌های نظری مرتبط بانفوذ در شبکه‌های اجتماعی، می‌توان به مولفه‌های مؤثر در محاسبه‌ی نفوذ دست پیدا کرد. در این مقاله با تعریف شاخص‌ها و معیارهای شبکه‌ای فعالیت عامل‌ها متناسب با عملیات نفوذ، نفوذ عامل‌ها محاسبه می‌شود. در روش پیشنهادی، ابتدا مدلی جهت ارزیابی عامل با توجه به ویژگی‌های بااهمیت برای ارزیابی عملیات نفوذ معرفی شده است و سپس با مجموعه داده‌های تولیدی متناسب، ارزیابی شده است. با توجه به شاخص‌های مورد استفاده در این مدل، مجموعه دادگانی که شامل همهٔ این شاخص‌ها باشد، وجود ندارد و ما سه مجموعه دادگان حاوی شاخص‌های مورد نظر از داده‌های توییتر تولید کردیم. نتایج به‌دست‌آمده نشانگر این مطلب است که مدل ارائه شده در کنار قابلیت تفسیرپذیری و عدم نیاز به داده‌های آموزشی، دارای عملکردی قابل مقایسه با روش‌های قبلی است.

کلیدواژه‌ها: شبکه‌های اجتماعی برخط، ارزیابی عملیات نفوذ، ارزیابی عامل‌های عملیات نفوذ

Keywords: Online social network (OSN), Influence operation assessment, influence operation agents assessment

[1] Hand-crafted features

* نویسنده مسئول: azgomi@iust.ac.ir

۱. مقدمه

دیکشنری آنلاین مریام-وبستر[1]، نفوذ[2] را «عمل و قدرت تولید اثر بدون اعمال آشکار نیرو و بدون به‌کارگیری دستور مستقیم» تعریف می‌کند. عملیات نفوذ نیز به‌عنوان یکی از انواع عملیات-های اطلاعاتی مطرح است[1]. عملیات نفوذ، در ده‌ها سال گذشته به‌عنوان ابزاری در دست کشورهای استعمارگر علیه جوامع کشورهای هدف اعمال شده است. با پدید آمدن و گسترش رسانه‌های جمعی مانند تلویزیون و رادیو، عملیات نفوذ شکل تازه‌ای به خود گرفت. اوج استفاده از رویکرد عملیات نفوذ از طریق رسانه‌های جمعی، در جنگ سرد میان شوروی و آمریکا صورت گرفت. در گام بعد، ظهور اینترنت و گسترش شبکه‌های اجتماعی مجازی، رویکرد بهره‌گیری از این فضا برای عملیات‌های نفوذ در دستور کار کشورهای بزرگ قرار گرفت. گسترش شتابان .و فراگیری رسانه‌های اجتماعی در فضای سایبری، امکان طرح-ریزی و اجرای عملیات‌های نفوذ فراگیر علیه مخاطبان هدف در هر کجا از دنیا را فراهم آورد. در یک دهۀ گذشته، بر اثر رشد سریع شبکه‌سازی اجتماعی برخط (مانند فیس‌بوک[3] و لینکدین[4]) و رسانۀ اجتماعی برخط (مانند توییتر و اینستاگرام)، نفوذ اجتماعی محاسباتی و به‌ویژه پویایی‌های انتشار[5] نفوذ اجتماعی، توجه تحقیقاتی زیادی را در حوزه‌های داده‌کاوی، یادگیری ماشین، نظریه بازی الگوریتمی و علوم کامپیوتر نظری به خود جلب کرده است[2].

با گسترش محبوبیت شبکه‌های اجتماعی برخط و رسانه‌های اجتماعی[6] در فضای سایبر، آنها به‌عنوان ابزاری راهبردی برای طرح‌ریزی، اجرا و ارزیابی عملیات نفوذ مطرح شده‌اند. اما نکته جالب آن است که کشورهای غربی به‌ویژه آمریکا از اعتراف صریح به به‌کارگیری این بستر برای عملیات نفوذ خودداری می‌کنند. بر اساس بررسی صورت گرفته، در هیچ‌کدام از اسناد منتشر شده توسط سازمان‌های غربی، از فضای سایبری به‌عنوان بستری برای اجرای عملیات نفوذ ذکر نشده است. البته برای اولین بار، در سال ۲۰۱۸، آمریکا طی گزارشی ایران را به انجام عملیات نفوذ در شبکه‌های اجتماعی برخط علیه اروپا و آمریکا متهم کرد[3].

بر اساس مفهوم و هدفی که برای عملیات نفوذ در نظر گرفته می‌شود، طرح‌ریزی و ارزیابی آن متفاوت خواهد بود. روش‌های متعدد کیفی و کمی برای سنجش و ارزیابی میزان موفقیت عملیات‌های نفوذ مطرح است. دو حالت کلی ارزیابی نفوذ قابل تصور است: یکی ارزیابی عملکرد یک گروه عملیاتی در انجام عملیات نفوذ است و دیگری ارزیابی اثر اجرای یک عملیات نفوذ است. موضوع مورد علاقه این تحقیق، ارزیابی عملکرد یک گروه عملیاتی عملیات نفوذ است. طبیعی است که هر گونه ارزیابی عملکرد گروه عملیاتی نفوذ، مبتنی است بر ارزیابی عملکرد عامل‌های اصلی این گروه عملیاتی است. با توجه به حجم انبوه پیام‌ها و کنش‌ها در هر عملیات نفوذ در رسانه‌های اجتماعی فضای سایبری، از روش‌های محاسباتی جهت ارزیابی عملکرد عامل‌های عملیات نفوذ، بهره گرفته می‌شود.

پرداختن به ارزیابی عملیات نفوذ با رویکرد محاسباتی مستلزم بهره‌گیری از رویکرد میان‌رشته‌ای است. همچنین مفهوم و شاخص‌های کیفی سنجش اثر عملیات نفوذ بر اساس دانش مربوطه در علوم روان‌شناسی اجتماعی و شناخت اجتماعی قابل برداشت است و از سویی نحوه استخراج و سنجش این شاخص‌ها بر مبنای داده‌های حجیم مربوط به بستر اجرای عملیات نفوذ، با بهره‌گیری از فناوری‌های محاسباتی حوزه مهندسی کامپیوتر و هوش مصنوعی مانند مدل‌سازی محاسباتی، یادگیری ماشین، تحلیل محاسباتی شبکه‌های اجتماعی حاصل می‌شود.

مدل‌های محاسباتی، ابزار مناسبی در دست تحلیل‌گران شبکه‌های اجتماعی برخط برای بهتر دیدن صحنه عملیات‌های نفوذ و ایجاد آگاهی وضعیتی دقیق‌تر و به‌روزتر است. به کمک رویکرد محاسباتی، می‌توان به پرسش کلیدی «شاخص‌های[7] ارزیابی کارکرد عملیات نفوذ در شبکه اجتماعی برخط مورد نظر و روش‌های سنجش آن‌ها کدام‌اند؟» پاسخ داد.

مسئله این تحقیق را می‌توان ذیل حوزه تحقیقی تحلیل محاسباتی نفوذ در شبکه‌های اجتماعی برخط دانست. مسائل مهم این حوزه شامل یافتن روشی محاسباتی برای حداکثرسازی نفوذ، یافتن گره با بیشترین نفوذ، یافتن الگوی نفوذ، تحلیل رقابتی نفوذ، حداقل‌سازی نفوذ گره مخرب، سنجش نفوذ جمعی یک انجمن، استخراج الگوی تکاملی نفوذ است.

تفاوتی که روش ارائه شده نسبت به روش‌های موجود دارد، در برداشتن تمام ویژگی‌های زیر به طور هم‌زمان است:

- امکان تحلیل مولفه‌های تأثیرگذار در ارزیابی افراد به‌صورت مجزا: با رویکردی که این روش دارد، امکان تحلیل کاربران از جنبه‌های مختلف در کنار ارزیابی کلی نفوذ آن‌ها فراهم می‌شود.

[1] Merriam-Webster
[2] influence
[3] Facebook
[4] LinkedIn
[5] propagation
[6] Social media
[7] Index

نویسنده مسئول: azgomi@iust.ac.ir

- وزن‌دهی متفاوت به کنش‌های مختلف کاربران: نیاز است بین اهمیت کنش‌ها از جمله پسند، پاسخ و ری‌توییت تفاوت قائل شد
- توجه به تنوع اجتماعی کاربران واکنش‌دهنده: به بیان دیگر، واکنش کاربری که از لحاظ اجتماعی در انجمنی مشابه با من قرار دارد با کاربری که در انجمنی متفاوت قرار گرفته است، متفاوت تلقی می‌شود.
- فاصلهٔ زمانی واکنش‌ها: بازه زمانی بین کنش یک کاربر و واکنش دیگر کاربران به آن، موضوعی است که اهمیت زیادی در سنجش نفوذ دارد.
- توجه به میزان نفوذپذیری کاربران: اگر میزان تأثیرپذیری فردی از دیگران بالا باشد، تأثیرگذاری بر او، نشانه‌ای از توان بالای فرد تأثیرگذار نیست، لذا نیاز است نفوذپذیری افراد مورد توجه قرار گیرد.
- توجه به شباهت کاربران با یکدیگر: فارغ از اینکه دو نفر از لحاظ اجتماعی در انجمن‌های مشابه باشند یا نباشند، داشتن شباهت زیاد یا کم در ارزیابی نفوذ یک نفر بر نفر دیگر تأثیرگذار است. این شباهت بر اساس ارتباطات دنبال‌کنندگی و ری‌توییت سنجیده می‌شود.

مسئله اصلی ما در این تحقیق، چگونگی مدل‌سازی محاسباتی نفوذ عامل عملیات نفوذ در شبکه‌های اجتماعی برخط جهت ارزیابی شبکه‌ای این عملیات است. در این راستا ابتدا با مفاهیمی همچون عملیات نفوذ، مدل‌سازی محاسباتی و شبکه‌های اجتماعی برخط آشنا خواهیم شد و سپس به بررسی کارهای انجام شده در این زمینه خواهیم پرداخت. در گام بعد، مدلی محاسباتی را تشریح خواهیم کرد. در نهایت نیز به انجام آزمایشاتی روی داده‌های دنیای واقعی خواهیم پرداخت تا مدل ارائه شده را ارزیابی نماییم.

۲. ادبیات موضوع

۲-۱. محیط اطلاعاتی و عملیات نفوذ

محیط اطلاعاتی شامل مجموعه‌ای از افراد، سازمان‌ها و سیستم‌هایی هست که جمع‌آوری، پردازش، انتشار یا اقدامی را بر روی اطلاعات انجام می‌دهند[4]. این محیط، شامل سه بعد وابسته به هم است که به طور پیوسته با افراد و سازمان‌ها و سیستم‌ها تعامل می‌کنند. این ابعاد با عناوین فیزیکی، اطلاعاتی و شناختی معروف هستند. بعد فیزیکی مرکب است از سامانه‌های فرماندهی و کنترل، تصمیم‌سازهای کلیدی و زیرساخت پشتیبانی‌کننده که افراد و سازمان‌ها را در ایجاد تأثیر قادر می‌سازند. بعد اطلاعاتی، مشخص می‌کند که اطلاعات کجا و چگونه، جمع‌آوری، پردازش، ذخیره‌سازی، منتشر و محافظت شود. بعد شناختی، شامل اذهان کسانی است که ارسال کردن، دریافت کردن، پاسخ دادن و اقدام بر روی اطلاعات را انجام می‌دهند. بعد شناختی به فرایندهای شناختی مانند پردازش اطلاعات، ادراک، داوری و تصمیم‌سازی افراد یا گروه‌ها مرتبط می‌شود. این عناصر تحت تأثیر عوامل زیادی مانند باورهای فردی و فرهنگی، هنجارها، آسیب‌پذیری‌ها، انگیزش‌ها، احساسات، تجارب، اخلاقیات، آموزش، سلامت روانی، هویت‌ها و ایدئولوژی‌ها قرار دارند. در نظر گرفتن این عواملِ تاثیرگذاری در یک محیط داده شده، برای فهمیدن اینکه عملیات چگونه بر ذهن تصمیم‌ساز بهترین تأثیر را گذاشت و تاثیرات مطلوب را ایجاد کرد، بسیار مهم است. به همین علت، این بُعد مهمترین مولفه‌ی محیط اطلاعاتی را تشکیل می‌دهد. در این مقاله ما به دنبال ارائه‌ی مدلی جهت ارزیابی بُعد اطلاعاتی عملیات نفوذ هستیم.

در اسناد عملیات اطلاعاتی، از عملیات نفوذ به‌عنوان مجموعه اقداماتی که برای تحت تأثیر قرار دادن مخاطب هدف، تعریف شده است. مخاطب هدف، فرد یا گروهی است که برای نفوذ انتخاب شده است. قابلیت‌های مربوط اطلاعاتی، ابزار، تکنیک‌ها یا فعالیت‌هایی هستند که هر سه بعد از محیط اطلاعاتی را تحت تأثیر قرار می‌دهند. آنها توانایی مخاطب هدف را در جمع‌آوری، پردازش یا انتشار اطلاعات قبل و بعد ازتصمیم‌گیری مورد تأثیر قرار می‌دهد. هدف نهایی عملیات نفوذ، ایجاد یک تغییر در تصمیم و یا کنش جامعه مخاطب هدف به وضعیت مورد نظر است. یک تغییر، زمانی رخ می‌دهد که یک یا چند اثر که قدرت ایجاد یک نتیجه اقناعی را دارند، با استفاده از اقدامات جنبشی یا غیرجنبشی یا ترکیبی از آن‌ها صورت گیرد. این تغییر باید به‌صورت اثرات قابل اندازه‌گیری در دو وضعیت نگرشی[1] و یا رفتاری[2] باشد. اثرات مطلوب شامل مواردی مانند منع و تأخیر اطلاعات، تغییر نگرش، طولانی‌شدن تصمیم‌سازی و تغییر در رفتار مخاطب هدف است.

رسانه‌های اجتماعی سایبری نقش کلیدی در جنگ اطلاعاتی و به‌ویژه عملیات نفوذ دارند. ظهور وسایل ارتباطاتی جدید و سبک‌های جدید تعامل اجتماعی مجازی، زمینه اقناع انبوه مردم[3] را تغییر داده است. این محیط جدید، نبرد ایده‌ها برای قلب‌ها و ذهن‌ها را خیلی پیچیده می‌کند[7]. رسانه‌های اجتماعی مجازی به‌عنوان یک ابزار جنگ اطلاعاتی نقش مهمی در مناقشات ایفا می‌کند. هزینه کم و دسترسی آسان به ابزارهای رسانه اجتماعی به‌عنوان یک نیرویی عمل می‌کند که با افزایش قابلیت‌های شبکه‌سازی و سازماندهی، فزاینده می‌شود. قابلیت انتشار سریع

[1] attitudinal
[2] behavioural
[3] mass persuasion

تصاویر و نظرات برای شکل دهی روایت عمومی[1]، رسانه اجتماعی برخط را به‌عنوان یک سلاح راهبردی در دست کشورهای استعمارگر، تروریست‌ها، گروههای شورشی، یا دولت‌های درگیر در مناقشه تبدیل کرده است. از طرف دیگر مجموعه داده حجیم تولید شده از کنش‌های شبکه‌ای عامل‌ها، امکان بهره‌گیری از مدل‌ها و روش‌های محاسباتی برای ارزیابی نفوذ عامل‌ها را برای ما فراهم آورده است.

گاهی اوقات عامل‌های بانفوذ، رهبران افکار[2]، نوآوران[3]، بااعتباران[4] یا بازیگران توانا[5] نامیده می‌شوند. علاوه بر این، دسته-بندی‌های دیگری برای عامل‌ها، با توجه به گسترش نفوذ وجود دارد. به‌عنوان‌مثال، بعضی از محققان، رهبران افکار، نفوذکننده‌ها و مباحثه‌کنندگان[6] را بر اساس نوع فعالیت و تأثیرگذاری متمایز می‌دانند[5]. می‌توان بین شروع‌کنندگان[7] (عامل‌های که موضوع جدید را شروع می کنند) و منتشرکنندگان[8] (عامل‌های که مسئول انتشار آن موضوع هستند) تمایز قائل شویم. همچنین، عامل‌ها بانفوذ می توانند بر اساس محتوا و قدرتشان تقسیم‌بندی شوند.

دیگر عامل‌ها مرتبط، افراد مشهور[9] هستند. کسانی که تا حدی دارای معیارهای متفاوتی نسبت به نفوذکننده‌ها هستند. عامل‌ها می‌توانند مطابق با محبوبیتشان به پخش‌کننده‌ها[10] یا عامل‌ها غیرفعال[11] (با دنبال‌کنندگان زیاد و دنبال‌شوندگان کم)، آشنایان[12] (تعداد مساوی دنبال‌کننده و دنبال شونده) و عامل‌ها خبیث[13] (دنبال‌کننده کم و دنبال شونده زیاد مثل اسپمرها و بات‌ها) تقسیم شوند[6].

۲-۲. مدل‌سازی محاسباتی

.یک مدل محاسباتی شامل متغیرهای متعددی است که سیستم مورد مطالعه را مشخص می‌کند. مدل‌سازی محاسباتی، استفاده از پردازنده برای شبیه‌سازی و مطالعه سیستم‌های پیچیده با استفاده از ریاضیات، فیزیک و علوم کامپیوتر و مبتنی بر داده‌های مربوطه است. یک مدل محاسباتی شامل متغیرهای متعددی است که بر اساس آن سیستم مورد مطالعه، مشخص می‌شود.

[1] public narrative
[2] Opinion Leaders
[3] Innovators
[4] Prestigious
[5] Authoritative Actor
[6] Discussers
[7] Inventors
[8] Spreaders
[9] Celebrities
[10] Broadcasters
[11] Passive Users
[12] Acquaintances
[13] Miscreants/Evangelists

شبیه‌سازی با تنظیم متغیرها به‌تنهایی یا ترکیبی و مشاهده نتایج انجام می‌شود. مدل‌سازی محاسباتی امکان اجرای متعدد آزمایش شبیه‌سازی شده را توسط کامپیوتر فراهم می‌آورد. شبیه‌سازی، با تنظیم این متغیرها و مشاهده چگونگی تأثیر تغییرات بر نتایج پیش‌بینی‌شده توسط مدل انجام می‌شود.

از آنجا که فرایند نفوذ اجتماعی یک عامل در شبکه اجتماعی برخط را می‌توان یک سیستم پیچیده در نظر گرفت، برای سنجش نفوذ یک عامل در شبکه‌های اجتماعی برخط به مدل‌سازی محاسباتی مبتنی بر داده‌های آن نیاز است. شبکه‌های اجتماعی برخط داده کافی برای این تحلیل محاسباتی را فراهم می‌آورد. بنابراین، مدل‌سازی محاسباتی، ابزار مناسبی در دست تحلیل‌گران شبکه‌های اجتماعی برخط برای بهتر دیدن صحنه عملیات‌های نفوذ و ایجاد آگاهی وضعیتی دقیق‌تر و به‌روزتر است.

۳. کارهای انجام شده

مفهوم نفوذ، بر اساس زمینه تحقیق در کارهای تحقیقاتی مربوطه، متفاوت است و طبیعی است که بر اساس مفهومی که برای نفوذ در نظر گرفته می‌شود، مؤلفه‌ها و شاخص‌های سنجش متفاوتی استخراج می‌شود. در این بخش، به بررسی این رویکردهای متفاوت می‌پردازیم.

شاخص‌های شبکه‌ای مورد استفاده در ارزیابی نفوذ به‌طورکلی به شاخص‌هایی که فقط مبتنی بر ساختار هستند و شاخص‌هایی که علاوه بر ساختار، معنا را هم مورد توجه قرار می‌دهند تقسیم می‌شوند. با نگاه ساختاری، شبکه را می‌توان به‌عنوان گرافی به شکل $G = (V, E)$ مدل کرد، که V مجموعه گره‌ها و E مجموعه یال‌ها است. بر اساس قرارداد، گره معادل موجودیت‌ها و اتصال‌ها معادل روابط بین آنها است. حال اگر این شبکه اجتماعی برخط به معنای شبکه ارتباطات کاربران و عوامل کنشگر در یک رسانه اجتماعی برخط باشد، گره‌ها معادل کنشگرها و یال‌ها معادل ارتباطات اجتماعی صورت گرفته بین آنها خواهد بود. و مفهوم نفوذ یک گره هم به معنای میزان نفوذ اجتماعی یک کاربر خواهد بود. البته مفهوم نفوذ هم یک خاصیت نسبی دارد و نفوذ اجتماعی یک عامل بر عامل دیگری را در نظر گرفته می‌شود. یعنی نفوذ اجتماعی یک اثر جهتی[14] از گره A به گره B است و با قدرت یال از A به B مرتبط است. البته با در نظر گرفتن یک مدل ریاضی و محاسباتی تجمیعی، می‌توان محاسبه میزان اثر تجمیعی یک گره در یک شبکه اجتماعی برخط را محاسبه کرد.

[14] Directional effect

از دید ساختاری، معیارهای پایهٔ سنجش نفوذ، به دو دستهٔ کلی معیارهای مبتنی بر یال و معیارهای مبتنی بر گره تقسیم‌بندی می‌شوند و اغلب معیارها به گره‌ها در شبکه مرتبط می‌شوند. معیارهایی همچون قدرت رابطه[1] و مابینی یال[2] جزو معیارهای مبتنی بر یال و معیارها مرکزیت مختلف نیز مبتنی بر گره هستند. از طرف دیگر روش‌هایی وجود دارد که با توجه به نوع ارتباطات بین عامل‌ها، زمان هر کنش و تحلیل محتوا به سنجش نفوذ عامل‌ها می‌پردازند. روش‌های مبتنی بر ساختار و محتوا، غالبا بر پایهٔ رتبه‌صفحه[3] [8] می‌باشند. روش‌های ارائه شده در این قسمت را می‌توان به ۳ دستهٔ کلی تقسیم‌بندی کرد:

- روش‌هایی که در محاسبهٔ نفوذ به روابط دنبال‌کنندگی بین عامل‌ها و عکس‌العمل‌هایی که به هم نشان می‌دهند، توجه می‌کنند.
- روش‌هایی که در محاسبهٔ نفوذ، خط سیر زمانی را هم مورد توجه قرار می‌دهند.
- روش‌هایی که در محاسبهٔ نفوذ، محتوای پیام‌ها را هم مورد توجه قرار می‌دهند.

در یکی از روش‌های سادهٔ ارائه شده، مقاله[9] روش TURank (Twitter User Rank) را پیشنهاد داده است که در آن عامل‌ها و توئیت‌ها در یک گراف توییت-عامل ارائه می‌شوند. گراف توییت-عامل شامل گره‌های عامل و توئیت‌ها و یال‌ها نشان‌دهنده دنبال‌کردن و روابط ری‌توییت است. در این روش، توییتی بانفوذ است که توسط بسیاری از عامل‌ها بانفوذ ری‌توییت شده است و عاملی بانفوذ است که توسط بسیاری از عامل‌ها بانفوذ دنبال شده باشد و توئیت‌های مفید زیادی قرار می‌دهد.در روش‌های ارائه شده در [10] و[11] (روش رتبه‌کاربر[4]) نیز بر اساس تعداد توییت‌های عامل و تعداد دنبال‌کنندگان و دنبال‌شوندگان او امتیاز توییت و دنبال‌کردن برای عامل محاسبه می‌شود. مقاله [12] نیز با تعریف شاخصی به نام بزرگی نفوذ[5] به محاسبه نفوذ عامل مبتنی بر تعداد دنبال‌کنندگان او و عکس العمل‌های آن‌ها به پست‌های او می‌پردازد. در مقاله‌ای دیگر [13]، روشی مبتنی بر رتبه صفحه به نام starrank ارائه شده است که اهمیت یک عامل را با توجه به همسایگی‌اش در نظر می‌گیرد. روش MultiRank نیز که در [14] ارائه شده مبتنی بر یک مقیاس پیچیده برگرفته از رتبه صفحه است. این روش قدم تصادفی[6] روی گرافی از ریتوییت‌ها، پاسخ‌ها و دو رابطه اضافی بنام بازمعرفی[7] و خواندن[8] را در نظر می‌گیرد. بازمعرفی توئیت‌هایی است که شبیه به توئیت‌های قبلی دیگر عامل‌ها است اما بدون آنکه به منبع اصلی اشاره کند. خواندن، اشاره به احتمال خواندن توئیت‌های منتشر شده توسط دیگر عامل‌ها است.

با پیشرفت کار و پیچیده‌تر شدن ویژگی‌های مورد توجه جهت محاسبه نفوذ، مقاله‌ی [15] روشی ارائه داده که در آن به عدم تأثیرپذیری[9] عامل توجه می‌کند. نفوذ یک عامل فقط به تأثیرگذاری روی مخاطبان بستگی ندارد بلکه بر عدم تأثیرپذیری آنها نیز بستگی دارد. مقاله [16] مدعی شده است که می‌تواند عامل‌ها بانفوذ بالقوه در شبکه را بر اساس سطح فعالیت، تمایل به ری‌توییت[10] و نفوذ دو طرفه[11] پیدا کند. بعضی عامل‌ها ممکن است چندین حساب جعلی داشته باشند. این حساب‌ها ممکن است با یکدیگر در تعامل باشند و بنابراین امتیاز نفوذ خودشان را بهبود بدهند. دسته‌ای از روش‌ها به دنبال حذف تأثیر عامل‌ها جعلی در محاسبهٔ نفوذ بوده‌اند[17][18]. در مقاله‌ی [19] مفهوم پتانسیل شبکه‌های اجتماعی[12] که مبتنی بر نسبت ری‌توییت و منشن (r_{rt}) و نسبت تعامل[13] (r_i) محاسبه می‌شود را معرفی می‌کند.[20] نیز در محاسبه امتیاز عامل[14] تعداد توییت‌ها، ریتوییت‌ها، پاسخ‌ها و منشن‌ها را در نظر گرفته است. مقاله [21] به مطالعه‌ی روانشناختی اجتماعی افراد مشهور پرداخته و بیان می‌دارد که به‌طورکلی، سلبریتی کسی است که شناخته شده[15]، دوست داشتنی[16] و قابل شناسایی[17] توسط مردم در جامعه است. در مقاله [22] مدلی به نام عمل-عکس‌العمل[18] پیشنهاد شده که بر اساس وفاداری و توجه است. مقیاس عمل، میزان وفاداری طرفداران به سلبریتی را اندازه‌گیری می‌کند. توجه[19] از نظر پاسخ‌ها، منشن‌ها و ریتوییت‌ها در توییتر اندازه‌گیری می‌شود و عکس‌العمل، صفات مربوط به مقدار توجهی که سلبریتی برای هر کنش دارد را اندازه‌گیری می‌کند. یکی از روش‌هایی که به‌خوبی به تفکیک ویژگی‌هایی عامل پرداخته است، روشی است که در [23] ارائه شده است. در این پژوهش یک روش برای اندازه‌گیری کیفیت توئیت‌ها ارائه شده است که بر اساس آن نفوذ نویسندگان

[1] Tie strength
[2] Edge Betweenness
[3] PageRank
[4] UserRank
[5] Magnitude of Influence (MOI)
[6] random walks
[7] Reintroduce
[8] Read
[9] Passivity
[10] Willingness to retweet
[11] Pairwise influence
[12] Social Networking Potential (SNP)
[13] Interactor Ratio
[14] User influence score
[15] Acquiantance
[16] Likeability
[17] Identification
[18] Action-Reaction
[19] Attention

تعیین می‌شود. کیفیت توئیت‌ها مطابق با درجه تمرکز موضوعی و رفتار ری‌توییت ارزیابی می‌شود. در این مدل، امتیاز توییت، امتیاز نویسنده و امتیاز خواننده به‌صورت اندازه کیفیت یک توییت، نفوذ یک نویسنده و تمرکز موضوع یک خواننده تعریف شده است. این مقاله تنها مقاله‌ای است که تعداد نقل قول شدن را مورد توجه قرار داده است.در مقاله‌ی [24]معیاری مبتنی بر دو مفهوم محتوا[1] و مکالمه[2] ارائه شده است. محتوا به تعداد توییت‌های منتشر شده‌ی عامل و ارتباطات حاصل از این توییت-ها بستگی دارد و مکالمه نیز به تعداد پاسخ‌ها و تعداد دنبال-کنندگانی که با عامل از طریق ری‌توییت و منشن ارتباط داشته-اند، دقت می‌کند. در مقاله [5] الگوریتم‌های InfRank، LeadRank و DiscussRank که نفوذگران، رهبران و بحث-کنندگان را به ترتیب مشخص می‌کنند، معرفی شده است.

اما در دستهٔ دوم روش‌های ارائه شده، الگوریتم‌ها به بُعد زمان نیز توجه داشته‌اند. مقاله [25] با تعریف مفهوم گسترش-پذیری[3] و ارائه‌ی روش SpreadRank مبتنی بر معیار فاصله زمانی[4] ریتوییت‌ها و مکان عامل‌ها در درخت ری‌توییت، کار می-کند. روش ارائه شده در مقاله‌ی [26] که ProfileRank نام دارد، بر مبنای این فرض است که محتوای مرتبط، محتوایی است که توسط عامل‌ها پرنفوذ تولید و منتشر می‌شود و عامل‌ها پرنفوذ کسانی هستند که محتوای مرتبط تولید می‌کنند. یکی دیگر از کارهایی که با توجه به روند زمانی و درخت انتشار به محاسبه نفوذ پرداخته است، مقاله [27] است. ایده اصلی مقاله این است که توییت کننده، نفوذ را از واسطه‌هایش به ارث می‌برد. یکی از دیگر ویژگی‌هایی که مبتنی بر زمان است و در محاسبهٔ نفوذ تأثیر می‌گذارد، میرایی تأثیرات با گذشت زمان است. در [28] نویسندگان با در نظر گرفتن تأثیر زمان، ضریبی[5] برای محاسبه تأثیر توییت پیشنهاد داده‌اند که شامل تأثیر ری‌توییت و پاسخ است و به بازه زمانی توییت و عکس‌العمل هایی که به آن نشان‌داده‌شده بستگی دارد.

در قسمت سوم نیز الگوریتم‌هایی جای می‌گیرند که به محتوا هم توجه داشته‌اند. در [29] روشی برای اندازه‌گیری نفوذ عامل‌ها به نام TwitterRank ارائه شده است. در این روش ابتدا، موضوعاتی که مورد علاقه توییت‌کننده‌ها است با تحلیل محتوای توئیت‌ها جدا می‌شوند. بر اساس موضوعات جدا شده، شبکه روابط موضوعی میان توییت‌کننده‌ها ایجاد می‌شود. سرانجام الگوریتم TwitterRank اجرا می‌شود. مقاله [30] با ترکیب دو مفهوم محبوبیت[6] و شایستگی[7] به دنبال رتبه‌بندی رهبران افکار است. میزان محبوبیت با توجه به تعداد دنبال‌کنندگان عامل محاسبه می‌شود و در محاسبه میزان شایستگی، تعداد موضوعاتی که عامل روی آن پست قرار می‌دهد، تعداد پست مرتبط با هر موضوع و میزان ارتباط هر پست با هر موضوع اهمیت دارند. در مقاله [31] توئیت‌ها به سه دسته تقسیم‌بندی می‌شوند: توئیت-های اصلی[8] (OT)، توئیت‌های محاوره‌ای[9](CT)، توئیت‌های تکراری[10](RT). متریک‌هایی جهت ایجاد مجموعه‌ای از ویژگی‌ها برای هر عامل استخراج و ترکیب می‌شوند. دو مفهوم سیگنال موضوعی و قدرت سیگنال ارائه شده است. مفهوم اول بیانگر این هست که نویسنده صرف نظر از انواع توئیت‌های پست شده چه مقدار با موضوع مورد نظر درگیر است. مفهوم قدرت سیگنال نیز اصالت توئیت‌های نویسنده را نشان می‌دهد و هر چقدر به یک نزدیکتر باشد، یعنی نویسنده، قوی‌تر است. بعضی مقیاس‌ها به‌جای در نظر گرفتن تمام متن، توجهشان را به هشتگ‌ها محدود کرده‌اند. مقاله [32] در روشی که با نام TRank ارائه داده است، به شناسایی عامل‌ها بانفوذ توییتر روی موضوعات خاص که با هشتگ مشخص شده‌اند، می‌پردازد. TRank عامل‌ها را در سه بعد با توجه به روابط دنبال‌شوندگی، ری‌توییت شدن و پسند شدن رتبه‌بندی می‌کند. در [33] علاوه بر در نظر گرفتن آماره-های توییتری از جمله ری‌توییت، پاسخ و منشن، به بررسی شباهت محتوایی عامل‌ها هم پرداخته شده است.

در برخی مقالات، ویژگی‌های محتوایی، ویژگی‌های مبتنی بر زمان و دیگر ویژگی‌های ارتباطی مورد توجه قرار گرفته است. در مرور جامعی که در سال ۲۰۲۱ انجام شده است [34] ، تنها ۴ روش [35]، [36]، [37] و [38] از این روش‌ها به‌عنوان روش‌هایی شناسایی شده‌اند که در سنجش نفوذ کاربر، ویژگی‌هایی محتوایی را هم دخیل می‌کنند. در مقاله [39]، فرض بر این است که عاملی بانفوذ است که توانایی انتشار اطلاعات به خوانندگان مؤثر را داشته باشد و یک روش جدید برای پیدا کردن نفوذگران با در نظر گرفتن ساختار لینک و ترتیب زمانی اتخاذ اطلاعات در توییتر ارائه شده است. در این مقاله تعداد کل عامل‌ها که یک توییت در رابطه با هر موضوع را دریافت کرده‌اند، محاسبه می-شود و بررسی می‌شود که چگونه این تعداد به مرور زمان تغییر می‌کند. مقاله [40] روشی برای اندازه‌گیری نفوذ عامل‌ها با در نظر گرفتن مخاطبان صریح و ضمنی و شمارش تعداد مخاطبان

[1] Content
[2] Conversation
[3] Spreadability
[4] Time interval
[5] Time-effectiveness Attenuation Coefficient (TAC)
[6] Popularity
[7] Competency
[8] Original tweet
[9] Conversational tweet
[10] Retweet

مؤثر[1] ارائه داده است. مدل نفوذ توییت، شامل نفوذ روی هر دو مخاطب صریح و ضمنی تحت تأثیر مسیرهای انتشار اطلاعات است. مخاطبان مؤثر ضمنی توییت، عامل‌های هستند که توییت را بدون پست گذاشتن یا پاسخ مجدد روی آن می‌خوانند. علاوه بر این فرض شده که وقتی یک عامل توییت را ری‌توییت می‌کند یا پاسخ می‌دهد، نشان دهنده این است که عامل توییت را خوانده است و یک مخاطب مؤثر صریح است. همچنین عامل‌های که به ریتوییت‌های یک توییت، پاسخ داده‌اند هم جزو مخاطبان مؤثر صریح هستند. بنابراین مخاطبان مؤثر صریح مخاطبانی هستند که توییت را به‌صورت مستقیم یا غیرمستقیم ری‌توییت کرده‌اند یا پاسخ داده‌اند. این کمک می‌کند تا بازیگران کلیدی در فرایند انتشار اطلاعات را شناسایی کنیم. هر دو ساختار شبکه دنبال‌کنندگی و تعامل میان عامل‌ها (عمل‌های پاسخ و ری‌توییت) برای اندازه‌گیری نفوذ یک توییت در نظر گرفته شده است. نفوذ توییت به‌صورت تعداد مخاطبانی که توییت را در طول دوره زمانی t خوانده‌اند نامیده می‌شود. علاوه بر این، به‌موقع بودن[2] یک فاکتور مهم برای اندازه‌گیری نفوذ عامل است. در این مقاله نفوذ عامل بر اساس نفوذ توییت در موضوعات و دوره‌های زمانی مختلف اندازه‌گیری می‌شود. در مقاله [41]، نفوذ یک عامل به وسیله کیفیت توئیت‌های پست شده[3] و فعال بودن[4] آن عامل سنجیده می‌شود. فعال بودن یک عامل توسط شاخص‌های مختلف تعیین می‌شود: تعداد توئیت‌ها، تعداد دنبال‌کننده‌ها و تعداد دنبال‌شونده‌ها. شاخص‌های ارزیابی فعال بودن عامل‌ها به دو دسته تقسیم می‌شوند. شاخص‌های فعال که تحت کنترل خود عامل است، مثل تعداد دنبال‌شونده‌ها و تعداد توئیت‌ها و شاخص‌های غیر فعال که تحت کنترل دیگر عامل‌ها هستند، مثل تعداد دنبال‌کننده‌ها. رفتار عامل‌ها نسبت به یک توییت به معنی کیفیت توییت است که با توجه به درجه تمرکز موضوع، رفتار ری‌توییت شدن و تأثیر موضوع خاص عامل‌های که آن را ری‌توییت می‌کنند، تعیین می‌شود. در مقاله [35]، روش جدیدی پیشنهاد شده است که ویژگی‌های متنی و غیر متنی را برای شناسایی متخصصان حوزه فناوری اطلاعات در یک شبکه میکروبلاگ ترکیب می‌کند. این مطالعه از رویکردهای حریصانه رو به جلو و عقب برای ترکیب بهینه‌ی ویژگی‌ها و الگوریتم SVM برای رتبه‌بندی متخصصان حوزه استفاده می‌کند. نتایج نشان می‌دهد که رویکرد حریصانه به عقب ترکیب بهینه‌ی ویژگی‌ها را تولید می‌کند. متعاقباً متخصصان حوزه موضوعی بر اساس ویژگی‌ها در ترکیب بهینه‌ی ویژگی‌ها رتبه‌بندی می‌شوند که منجر به دقت بالاتر می‌شود. در مقالهٔ [36]، روشی پیشنهاد شده است که ویژگی‌های اطلاعات و کیفیت کاربر را برای اندازه‌گیری مؤثر نفوذ پیام در مقایسه با مطالعات موجود ترکیب می‌کند. علاوه بر این، نفوذ ارسال کنندگان هرزنامه را می توان با استفاده از ضریب مجازات کاهش داد. ویژگی‌های اطلاعاتی مانند ری‌توییت و پاسخ نشان‌دهنده تأثیرات متفاوتی هستند. کاربران ری‌توییت‌کننده نقش فعالی در انتشار اطلاعات دارند در حالی که کاربران پاسخ‌دهنده بیشتر به‌عنوان مشارکت کنندگان در تقویت انتشار اطلاعات عمل می کنند. بنابراین، این مطالعه وزن‌های مختلفی را برای ریتوییت‌ها و پاسخ‌ها در نظر می‌گیرد و به‌صورت پویا تأثیر پیام را اندازه‌گیری می‌کند. متعاقباً، این مطالعه، کیفیت آن دسته از کاربرانی را که در فرایند انتشار پیام شرکت می‌کنند را با توجه به ترتیب زمانی در نظر می‌گیرد. در میکروبلاگ‌ها، پیامی بانفوذ بالا بر برخی از کاربران باکیفیت بالا تأثیر می‌گذارد. یک کاربر باکیفیت بالا دنبال‌کنندگان زیادی دارد و دنبال‌کنندگان کمی دارد. در مرحله بعد، کاربرانی که کیفیت کاربری زیر یک آستانه مشخص دارند با استفاده از یک تابع مجازات می‌شوند. در نهایت، تأثیر پیام با استفاده از تعداد ری‌توییت، نظرات و کیفیت کاربر اندازه‌گیری می‌شود. علاوه بر این، مدل به‌صورت دینامیکی کاهش تأثیر در فرایند انتشار پیام را محاسبه می‌کند. کیفیت کاوی معنایی یک رویکرد مبتنی بر کیفیت مهم است [37]. برخلاف مطالعات موجود، کاوش معنایی، امتیاز نفوذ را با استفاده از ترکیبی از مرکزیت شبکه و کیفیت متنی توییت ها محاسبه می‌کند. این فرایند از اعتبارسنجی و تجزیه توییت‌ها از مجموعه داده‌ها و استخراج ویژگی‌های مرتبط برای محاسبه امتیاز نفوذ هر کاربر شروع می‌شود. ویژگی‌های استخراج‌شده بعداً برای تولید یک گراف اتصال، استفاده می‌شوند که روابط دنبال‌کنندگی را نشان می‌دهد. سپس گراف اتصال برای شناسایی جوامع با استفاده از روش Louvain تجزیه و تحلیل می‌شود. پس از آن، دو نوع تحلیل مختلف برای شناسایی کاربران تأثیرگذار انجام می‌شود. تجزیه و تحلیل شبکه، امتیاز نفوذ کاربران شبکه را با استفاده از ویژگی‌های استخراج شده محاسبه می‌کند. از سوی دیگر، تجزیه و تحلیل کیفیت بر اساس متریک Kincaidi برای محاسبه میانگین تعداد هجاها در هر کلمه و همچنین میانگین طول جمله است. معیار Flesch–Kincaid که به طور خاص برای انگلیسی طراحی شده است، در واقع یک تست خوانایی برای قضاوت در مورد سطح دشواری یک متن است. تجزیه و تحلیل کیفیت با ترکیب نمره Kincaidi با تعداد کلمات و تجزیه و تحلیل متن دیل-چال که از طول کلمه برای تعیین دشواری یک کلمه برای خواننده استفاده می‌کند، بهبود می‌یابد. در نهایت، کاربران بر اساس شبکه تجمیعی و

[1] Effective audience
[2] Timeliness
[3] Quality score
[4] Activity score

تجزیه و تحلیل کیفیت رتبه‌بندی می‌شوند. یک دسته از روش‌ها به بررسی نفوذ کاربران در رویدادهای خاص می‌پردازند. اکثر روش‌های اصلی موجود از ویژگی‌های نسبتاً ساده، ارتباطات رفتاری کاربران یا ارتباطات محتوایی برای ساخت شبکه استفاده می‌کنند. علاوه بر این، این روش‌ها قادر به در نظر گرفتن ویژگی‌های کاربر یا بررسی ویژگی‌های مختص رویداد نیستند. مطالعه [38]، یک رویکرد محاسباتی چند جانبه جهت محاسبه اهمیت کاربر در رویدادهای مشخص، پیشنهاد کرده است. یک کاربر در شبکه به ترتیب در چهار سطح در کاربر، لایه طرفداران، لایه میکروبلاگ و لایه رویداد ارزیابی می‌شود. نویسندگان اهمیت لایه‌های مختلف کاربر را با بررسی ویژگی‌های مهم کاربران در هر سطح و همچنین تأثیر این ویژگی‌ها بر اهمیت کاربر مربوطه محاسبه می‌کنند. در نهایت، این مطالعه اهمیت کاربر را در سطوح مختلف برای محاسبه اهمیت کلی در رویداد یکپارچه می‌کند. در

جدول ۱ ویژگی‌های مقالات بررسی شده با رویکرد ساختاری-محتوایی، مشخص شده است.

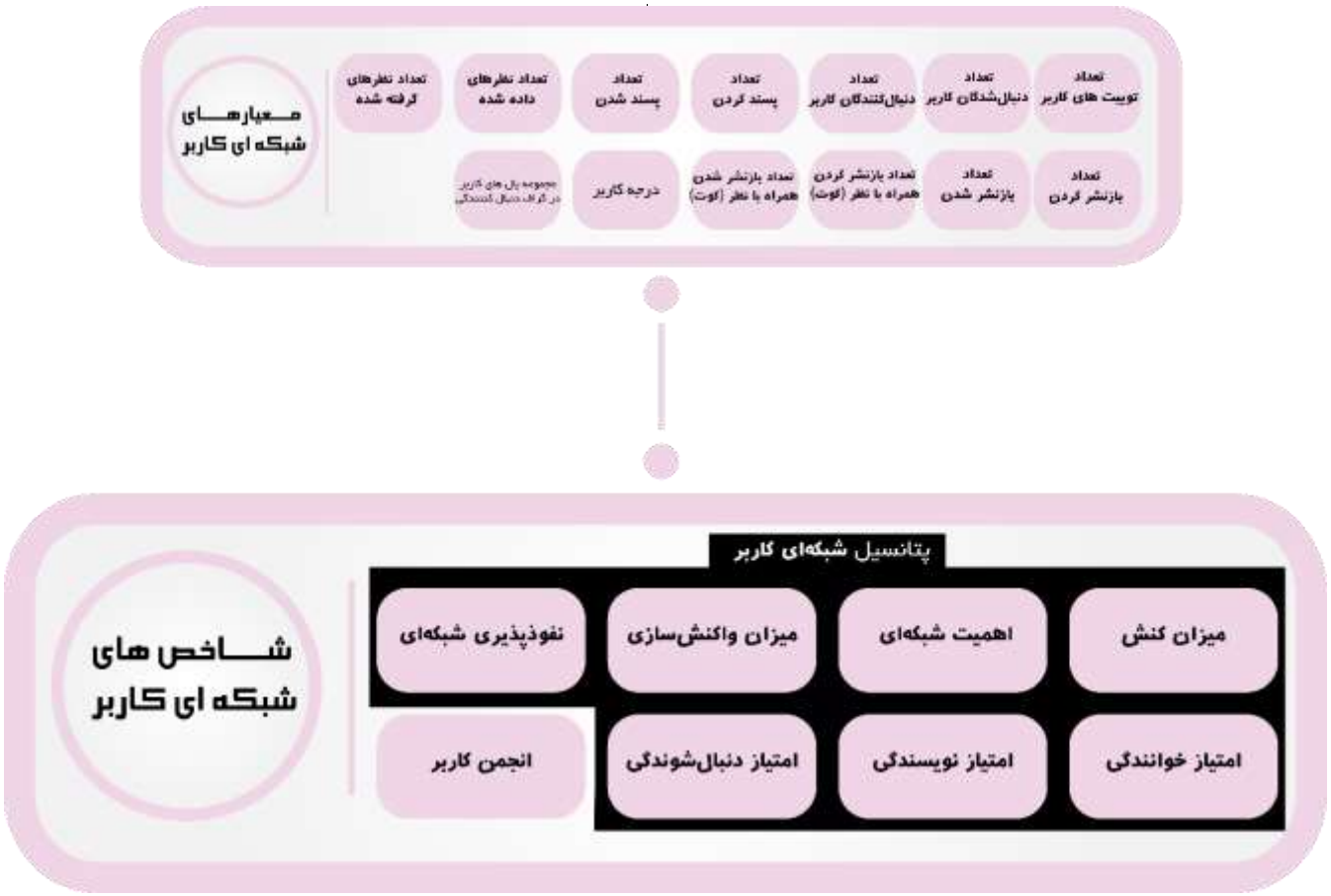
جدول ۱: ویژگی‌های مقالات حوزه‌ی سنجش نفوذ عامل با رویکرد ساختاری محتوایی

نوع روش	شماره مرجع	سال	دنبال کردن/ شدن	توییت	ری‌توییت	پاسخ	منشن	پسند	روند زمانی	محتوا	مفاهیم و ایده‌های ارائه شده
مبتنی بر محتوا	[38]	2020	✓	✓	✓	✓		✓		✓	ارزیابی کاربر در ۴ سطح کاربر، طرفداران، میکروبلاگ و رویداد
	[37]	2017	✓	✓	✓		✓		✓	✓	سنجش کیفیت متن از لحاظ زبان‌شناسی
	[36]	2016	✓	✓	✓	✓	✓		✓	✓	حذف اثر هرزنامه‌ها، توجه به مدل و زمان انتشار پیام، تفاوت قائل شدن بین تأثیر ری‌توییت و پاسخ، توجه به کیفیت افرادی که عکس‌العمل به پیام نشان داده‌اند
	[35]	2015	✓	✓	✓	✓	✓	✓		✓	انتخاب ویژگی‌های بهینه از بین تمام ویژگی‌ها
	[41]	2013	✓	✓	✓				✓	✓	کیفیت توییت و فعال بودن
	[40]	2013	✓		✓	✓			✓	✓	مخاطبان صریح، ضمنی و مؤثر، به‌موقع بودن
	[39]	2010	✓						✓	✓	خوانندگان مؤثر
	[33]	2013	✓		✓	✓	✓			✓	شباهت معنایی
	[32]	2015	✓		✓			✓		✓	نفوذ دنبال شدن، ری‌توییت شدن و پسند شدن
		2011		✓	✓	✓				✓	سیگنال موضوعی و قدرت سیگنال، اصالت توییت
	[30]	2016	✓	✓						✓	رهبران نظر،محبوبیت و شایستگی
	[29]	2010	✓							✓	شبکه روابط موضوعی
مبتنی بر زمان	[41]	2013	✓	✓	✓				✓	✓	کیفیت توییت و فعال بودن
	[40]	2013	✓		✓	✓			✓	✓	مخاطبان صریح، ضمنی و مؤثر، به‌موقع بودن
	[28]	2013	✓	✓	✓	✓			✓		ضریب میرایی
	[27]	2013	✓		✓				✓		ارث‌بری نفوذ
	[26]	2013		✓	✓				✓		ارتباط بین علائق مخاطب و اطلاعات ارائه شده
	[39]	2010	✓						✓	✓	خوانندگان مؤثر
	[25]	2013			✓				✓		گسترش‌پذیری
دیگر ویژگی‌ها	[5]	2012	✓	✓	✓	✓	✓				نفوذگران، لیدرها، بحث‌کنندگان
	[24]	2011	✓	✓	✓	✓	✓				دو مفهوم محتوا و مکالمه

نوع روش	شماره مرجع	سال	دنبال کردن/ شدن	توییت	ری‌توییت	پاسخ	منشن	پسند	روند زمانی	محتوا	مفاهیم و ایده‌های ارائه شده
	[23]	2011	✓	✓	✓	✓					تمرکز موضوعی، امتیاز توییت، امتیاز نویسنده و امتیاز خواننده
	[22]	2014	✓		✓	✓	✓				مدل عمل عکس‌العمل، وفاداری و توجه
	[20]	2013		✓	✓	✓	✓				---
	[21]	2013	✓		✓	✓	✓				۳ ویژگی سلبریتی‌ها
	[12]	2011	✓		✓	✓		✓			تحلیل رفتاری، شاخص بزرگی نفوذ، تأثیر بر رفتار آینده مخاطبان
	[19]	2011	✓	✓	✓		✓				پتانسیل شبکه اجتماعی
	[18]	2015			✓	✓	✓				مقاوم در مقابل حملات سیبل
	[17]	2013	✓								تنبیه اسپم‌ها، تعداد بازیگران متقابل
	[16]	2012		✓	✓						سطح فعالیت، تمایل به ری‌توییت، عامل‌ها بانفوذ بالقوه
	[15]	2011	✓		✓						عدم تأثیرپذیری
	[14]	2013		✓	✓	✓					استفاده از walk random
	[13]	2010				✓	✓				starrank مثل رتبه صفحه
	[11]	2012	✓	✓							---
	[10]	2010	✓	✓							---
	[9]	2010	✓		✓						---
	[8]	2011									رتبه صفحه

۴. ارائه مدل محاسباتی

در این بخش، به ارائه مدل ارزیابی نفوذ عامل که مبتنی بر کمیت‌های شبکه‌ای است، می‌پردازیم. عامل‌ها دارای دو نوع کمیت با عناوین معیار و شاخص هستند. معیارهای عامل، کمیت‌های پایه‌ای مستخرج از داده‌های شبکه‌های اجتماعی برخط هستند. شاخص‌های عامل نیز مبتنی بر تحلیل معیارها و فرمول‌بندی آن‌ها برای یک مفهوم شبکه‌ای تولید می‌شوند.



شکل ۱ : معیارها و شاخص‌های تأثیرگذار در محاسبه‌ی توان شبکه‌ای انجمن

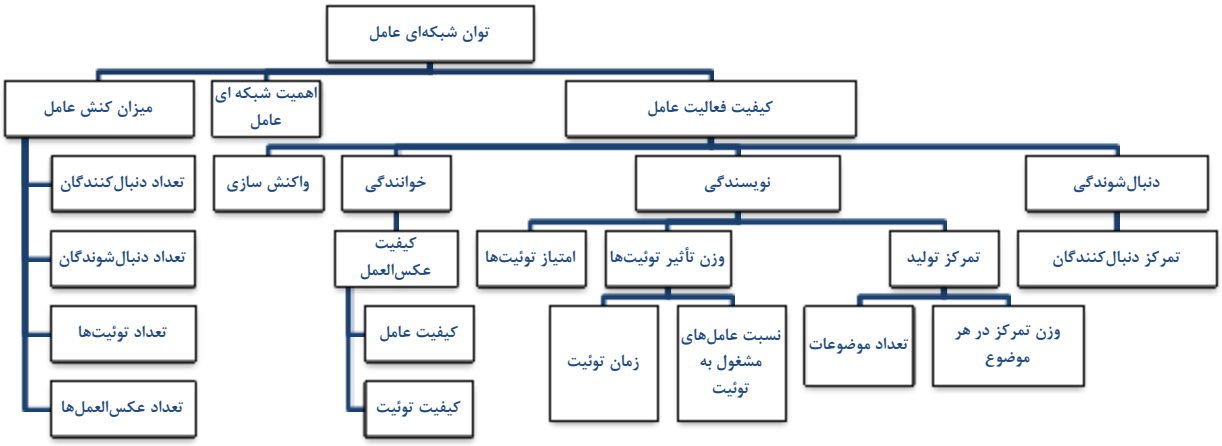
۴-۱. معیارهای شبکه‌ای عامل

بر اساس تعاملات عامل در شبکه‌های اجتماعی برخط می‌توان فهرستی از معیارهای شبکه‌ای عامل به شرح زیر ارائه داد:

- تعداد توییت‌های عامل ($|Tw(u)|$)
- تعداد دنبال‌شدگان عامل ($|Fl(u)|$)
- تعداد دنبال‌کنندگان عامل ($|Fl_x(u)|$)
- پسندهایی که عامل انجام داده است ($|Like(u)|$)
- پسندهایی که به عامل داده شده است ($|Like_x(u)|$)
- نظراتی که عامل گذاشته است ($|Cm(u)|$)
- نظراتی که به عامل داده شده است ($|Cm_x(u)|$)
- بازنشرهایی (ری‌توییت) که عامل انجام داده است ($|Rt(u)|$)
- بازنشرهایی (ری‌توییت) که از عامل انجام شده است ($|Rt_x(u)|$)
- نقل‌قول‌هایی که عامل انجام داده است ($|Qt(u)|$)
- نقل‌قول‌هایی که به عامل داده شده است ($|Qt_x(u)|$)
- درجه عامل ($D(u)$)
- مجموعه یال‌های عامل ($E(u)$)

۴-۲. شاخص‌های شبکه‌ای عامل

شاخص‌های عامل به‌طورکلی به دو دسته تقسیم می‌شوند که یکی از آن‌ها انجمن (طیف) عامل است و بقیه شاخص‌ها در کنار هم، عنوان توان شبکه‌ای می‌گیرند. این شاخص‌ها که در شکل ۱ به نمایش گذاشته شده‌اند، در ادامه شرح داده شده‌اند.



شکل ۲: مولفه‌های تأثیرگذار در توان شبکه‌ای عامل

- توان شبکه‌ای عامل: توان شبکه‌ای یک عامل به سه مؤلفه میزان کنش عامل، اهمیت شبکه‌ای عامل و کیفیت فعالیت عامل به ترتیب زیر وابسته است (این مولفه‌ها در شکل ۲ نمایش داده شده‌اند):

$$UserPotential(u) = \alpha_p Activity(u) + \beta_p Importance(u) + \lambda_p Quality(u) \quad (۱)$$

در رابطه بالا، با فرض اینکه مجموع ضرایب α_p، β_p و λ_p برابر یک است، این ضرایب به‌صورت زیر محاسبه می‌شوند [42]:

۱. نرمال‌سازی داده‌های کنش، اهمیت شبکه‌ای و کیفیت فعالیت به‌صورت جداگانه

۲. محاسبه آنتروپی داده‌های نرمال‌سازی شدهٔ کنش، اهمیت شبکه‌ای و کیفیت فعالیت به‌صورت جداگانه (نحوه محاسبه آنتروپی برای کنش نمایش داده شده است.)

$$E_{Activity} = -\frac{1}{\ln(n)} \sum_{i=1}^{n} Activity(u).\ln\big(Activity(u)\big) \quad (۲)$$

۳. محاسبهٔ ضریب هر فاکتور (نحوه‌ی محاسبه برای کنش نمایش داده شده است.)

$$\alpha_p = \frac{1-E_{Activity}}{\sum_{j\in\{Activity.Importance.Quality\}}(1-E_j)} \quad (۳)$$

نحوه‌ی محاسبه‌ی همهٔ ضرایبی که در بخش‌های دیگر فرمول نیز به کار برده شده‌اند، به همین صورت خواهد بود.

- حجم کنش عامل: این کمیت، تابعی از تعداد دنبال‌کنندگان، دنبال‌شوندگان، توییت‌ها و عکس‌العمل‌هایی (لایک، پاسخ، ری‌توییت و نقل قول) است که عامل آن‌ها را انجام داده است و به‌صورت زیر محاسبه می‌شود.

$$Activity(u) = \alpha_{A1}\frac{|Fl(u)|}{T_{end}-T_{first}} + \alpha_{A2}\frac{|Tw(u)|}{T_{end}-T_{first}} + \alpha_{A3}\log|Fl_x(u)| + \alpha_{A4}\frac{|Cm(u)|}{T_{end}-T_{first}} + \alpha_{A5}\frac{|Rt(u)|}{T_{end}-T_{first}} + \alpha_{A6}\frac{|Qt(u)|}{T_{end}-T_{first}} \quad (۴)$$

زمانی که به دنبال محاسبه‌ی نفوذ در یک موضوع خاص هستیم، بازهٔ رخداد آن موضوع مشخص‌کنندهٔ زمان شروع (T_{first}) و پایان (T_{end}) است.

- اهمیت شبکه‌ای عامل: اهمیت شبکه‌ای عامل را با استفاده از مقدار page rank عامل نمایش می‌دهیم.

$$Importance(u) = PageRank(u) \quad (۵)$$

- کیفیت فعالیت عامل: این شاخص مبتنی بر چهار ویژگی واکنش‌سازی، خوانندگی، نویسندگی و دنبال‌شوندگی عامل محاسبه می‌شود.

$$Quality(u) = \alpha_q RM(u) + \beta_q ReadingScore(u) + \lambda_q AuthorScore(u) + \gamma_q FollowshipScore(u) \quad (۶)$$

  - واکنش‌سازی عامل: تعداد عکس‌العمل‌های نشان‌داده‌شده به عامل در واحد زمان، در یک بازهٔ زمانی مشخص، نشانگر میزان واکنش‌سازی عامل است.

$$RM(u) = \frac{1}{T_{end}-T_{first}} * \begin{pmatrix} |Like_x(u)| + |Cm_x(u)| \\ +|Rt_x(u)| + |Qt_x(u)| \end{pmatrix} \quad (۷)$$

  - امتیاز خوانندگی عامل: این معیار در حقیقت بیانگر کیفیت عکس‌العمل‌هایی است که عامل از خود نشان داده است. عکس‌العمل‌ها چهار نوع لایک، پاسخ، ری‌توییت و نقل قول هستند. این کیفیت با توجه به کیفیت مخاطبان عامل و کیفیت پیام‌هایی که عامل به آن‌ها عکس‌العمل نشان داده است، محاسبه می‌شود. هر چقدر عامل‌هایی که عامل مورد نظر به آن‌ها عکس‌العمل نشان داده، کیفیت بیشتری داشته

باشند، نشان از این دارد که عاملی مورد نظر در انتخاب عامل‌هایی که مشاهده می‌کند، بادقت و هوشمندی بیشتری عمل کرده است. از طرف دیگر، هر چقدر کیفیت پیام‌هایی که عامل مورد نظر به آن‌ها عکس‌العمل نشان داده است بالاتر باشد، به معنای کیفیت بالای این عامل در خوانندگی دارد.

$$ReadingScore(u) = \frac{1}{|Engagements|} \times \left(\text{کیفیت مخاطبان عامل} + \text{کیفیت پیام هایی که عامل به آن ها مشغولیت نشان داده است}\right) = \frac{1}{|Engagements|} \times \left(\sum_{e=1:4}\left(\sum_{k\in\{u(engaged\ users)\}}|Engagement(e)| \times UserScore(k)\right) + \left(\sum_{m\in engaged\ massages}|Engagement(e)| \times MessageScore(m.k)\right)\right) \quad (۸)$$

با توجه به تکراری بودن فرایند محاسبه نفوذ پیام و عامل، در محاسبه‌ی نفوذ عامل در تکرار t، مقدار نفوذ پیام در تکرار t-1 مورد استفاده قرار می‌گیرد و بالعکس.

مقدار اولیه امتیاز پیام به‌صورت زیر محاسبه می‌شود:

$$MessageScore(m_u)_{t_0} = \frac{1}{2.5} \times \left(\left(\text{تعداد پسند} \times 0.25\right) + \left(\text{تعداد پاسخ} \times 0.75\right) + \left(\text{تعداد ری‌توییت} \times 0.5\right) + \left(\text{تعداد نقل قول} \times 1\right)\right) \quad (۹)$$

سپس امتیاز اولیه عامل برابر با مجموع امتیاز محتوای تولیدی توسط او قرار داده می‌شود.

- امتیاز نویسندگی عامل: امتیاز محتوای تولیدی، وزن تأثیر هر یک از محتواها و تمرکز موضوعی تولید محتوا، ۳ عامل تأثیرگذار در محاسبه‌ی امتیاز نویسندگی عامل هستند.

$$AuthorScore(u) = \gamma\left(\sum_{m\in(u's\ Messages)} w_{m\to u} \times MessageScore(m)\right) + \lambda Focus(u.T) \quad (۱۰)$$

$w_{m\to u}$ وزن تأثیر هر یک از پیام‌های عامل در افزایش امتیاز او را نشان می‌دهد. در محاسبه وزن تأثیر پیام در افزایش امتیاز عامل دو عامل نقش بازی می کنند: زمان پیام و تعداد عامل‌های مشغول شده به این پیام به نسبت تعداد کل عامل‌های فعال در موضوع. با در نظر گرفتن زمان پیام، در حقیقت به میرایی تأثیر پیام با گذشت زمان توجه کرده‌ایم. با توجه به آنچه در فرمول محاسبه‌ی EAC می‌بینید، پیام‌هایی که عامل جدیدا گذاشته باشد، تأثیر بیشتری در محاسبه‌ی نفوذ عامل خواهند داشت. این ضریب بین عامل‌هایی که قبلا فعالیت خوبی داشته‌اند ولی الان مانند قبل کار نمی‌کنند، با عامل‌هایی که اخیرا فعالیت خوبی داشته‌اند، حتی اگر حجم فعالیت هر دو عامل با هم برابر باشد، تفاوت قائل می‌شود و به عامل‌هایی که اخیرا فعالیت داشته‌اند، امتیاز بیشتری می‌دهد. می‌توان این ضریب را به این شکل نیز تفسیر کرد که هرچقدر بازه زمانی‌ای که عامل‌ها در آن مشغول به این پیام شده‌اند کمتر باشد ارزش مشغولیت[1] ایجاد شده بیشتر است.

$$EAC = \frac{1}{1+2^{time_{now}-CreationTime(m)}} \quad (۱۱)$$

با توجه به توضیحات بالا و دو مؤلفهٔ یاد شده، وزن تأثیر هر پیام در افزایش امتیاز عامل به‌صورت زیر محاسبه می‌شود.

$$w_{m\to u} = EAC * \frac{|\text{عامل‌ها مشغول روی این پیام}|}{|\text{کل عامل‌ها مرتبط با موضوع پیام}|} \quad (۱۲)$$

$Focus(u,T)$ نیز بیانگر میزان تمرکز موضوعی عامل است. تمرکز موضوعی نیز یکی از مواردیست که می‌تواند در بالا بردن کیفیت نویسندگی عامل مؤثر باشد.

$$Focus(u.T) = \frac{|\text{پیام های عامل } u \text{ در موضوع } T|}{|\text{کل پیام های عامل } u|} \quad (۱۳)$$

- امتیاز دنبال شوندگی عامل: تعداد دنبال‌کنندگان، امتیاز عامل‌ها دنبال‌کننده و تمرکز آن‌ها در دنبال‌کنندگی عوامل مورد توجه در محاسبه‌ی امتیاز دنبال‌شوندگی عامل هستند.

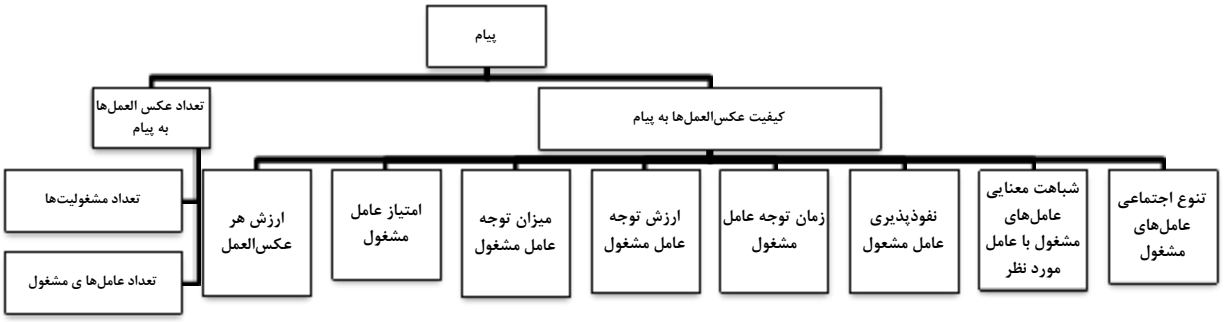
$$FollowshipScore(u) = \sum_{v\in follower(u)} \frac{UserScore(v)}{|followees(v)|} \quad (۱۴)$$

- انجمن عامل: تشخیص انجمن عامل با به‌کارگیری الگوریتم تشخیص انجمن مبتنی بر ماژولاریتی صورت می‌گیرد.

از طرف دیگر، امتیاز پیام به دو عامل مهم بستگی دارد. مورد اول تعداد عکس‌العمل‌ها به پیام و مورد دوم کیفیت عکس‌العمل‌هایی است که به پیام نشان‌داده‌شده است. معیارهای

[1] Engagement

دخیل در ارزیابی امتیاز پیام در شکل ۳ نمایش داده شده‌اند.



شکل ۳: عوامل تأثیرگذار در نفوذ پیام

با توجه به شکل ۳، می‌توان امتیاز یک پیام را به‌صورت زیر محاسبه نمود:

$$EngagementScore(m_u) = \left(\sum_{k\in(Engaged\ Users)}\sum_{e=1:3} value_e \times UserScore(k) \times Attention_e(k) \times AttentionValue_e(m) \times f(\Delta t) \times Passivity_e(k.u) \times \frac{1}{Similarity(u.k)} \times SD_e(u)\right) \quad (۱۵)$$

m_u نشانگر این است که تولیدکننده‌ی پیام m، عامل u بوده است. جهت اندازه‌گیری کیفیت عکس‌العمل‌های نشان‌داده‌شده به پیام، MessageScore محاسبه می‌شود. تعداد مشغولیت‌ها، تعداد عامل‌ها مشغول، ارزش هر عکس‌العمل، امتیاز عاملی که عکس‌العمل نشان داده است، میزان توجه عامل مشغول به پیام، ارزش توجه عامل مشغول برای پیام، زمان توجه عامل مشغول، نفوذپذیری عامل مشغول، شباهت معنایی عامل‌ها مشغول با عامل تولید کننده‌ی پیام و تنوع اجتماعی عامل‌ها مشغول، مواردی هستند که در محاسبه‌ی کیفیت عکس‌العمل‌ها، مورد توجه قرار می‌گیرند.

با توجه به نوع مشغولیت، می‌توان برای آن ارزشی تعریف کرد. ارزش عکس‌العمل‌ها ($value_e$) را به ترتیب برای ری‌توییت، پاسخ و نقل قول برابر با ۰,۵، ۰,۷۵ و ۱ در نظر گرفت.

میزان توجه عامل k به پیام m ($Attention_e(k)$)، بیانگر این است که میزان توجه عامل به این پیام، به نسبت توجهی که به کل پیام‌ها داشته چقدر است.

$$Attention_e(k) = \frac{1}{|\text{مشغولیت های عامل } k \text{ روی کل پیام ها در بازه زمانی } \Delta t|} \quad (۱۶)$$

ارزش توجه عامل k به پیام m ($AttentionValue_e(m)$) بیانگر میزان توجه عامل k به این پیام، به نسبت کل توجهی که توسط تمام عامل‌ها به این پیام شده است، است.

$$AttentionValue_e(m) = \frac{1}{|\text{مشغولیتهای کل عامل‌ها روی پیام } m \text{ در بازه زمانی } \Delta t|} \quad (۱۷)$$

زمان توجه عامل به پیام ($f(\Delta T)$) نیز به بررسی میزان طولانی بودن بازه‌ی زمانی بین تولید پیام و نشان‌دادن عکس‌العمل می‌پردازد.

$$f(\Delta T) = \frac{1}{\lambda} e^{-\frac{\Delta T}{\lambda}} \quad (۱۸)$$

با توجه به تابع بالا، هر چقدر فاصلهٔ عکس‌العمل نشان‌دادن به توییت کمتر باشد، یعنی نفوذی که توییت روی او گذاشته بیشتر بوده است. با تغییر مقادیر مختلف λ، اثر طولانی‌شدن زمان روی میزان نفوذ قابل تغییر است. در آزمایشاتی که ما انجام داده‌ایم (در بخش "۵.ارزیابی مدل ارائه شده")، λ را برابر یک قرار داده‌ایم. در این صورت میزان نفوذ به‌صورتی است که در جدول ۲نمایش داده شده است:

جدول ۲: تأثیر زمان عکس‌العمل با مقدار توجه عامل به پیام

$f(\Delta T)$	اختلاف زمانی بین عکس‌العمل و تولید پیام (به ساعت)
۱	۰
۰/۶۰	۰/۵
۰/۳۶	۱
۰/۱۳	۲

با توجه به نوع مشغولیت میزان نفوذپذیری عامل‌ها از هم

متفاوت است؛ مثلاً ممکن است نفوذپذیری عاملی در لایک کردن بالا باشد ولی در نقل قول کردن این‌طور نباشد؛ یا عاملی به لایک پیام‌های عاملی مخالف خود نپردازد ولی پاسخ‌هایی هم راستا با پیام عاملی مخالف خود بگذارد. در تمام فرمول‌های زیر پسوند e به معنی مختص بودن به نوع مشغولیت است. این معیار با توجه به فرمولی که در [۱۴] ارائه شده است، محاسبه می‌شود.

$$Relation_e(k.u) = \frac{|u's\ tweets\ that\ engaged\ k|}{|all\ u's\ tweet|} \quad (۱۹)$$

$$AcceptanceRate_e(k,u) = \frac{Relation_e(k,u)}{\sum_{l\in(users\ k\ is\ engaged\ on\ them)} Relation_e(k,l)} \quad (۲۰)$$

$$RejectionRate_e(k,u) = \frac{1-Relation_e(k,u)}{\sum_{l\in(users\ k\ is\ engaged\ on\ them)} 1-Relation_e(k,l)} \quad (۲۱)$$

$$Passivity_e(k,u) = RejectionRate_e(k,u) \times UserScore(u) \quad (۲۲)$$

نفوذپذیری عامل k از عامل u برابر با معکوس $passivity(k,u)$ است.

یکی از مؤلفه‌های تأثیرگذار در کیفیت مشغولیتِ گرفته شده، شباهت عامل‌ها مشغول شده با عامل مورد بررسی (Similarity(u,k)) است. در واقع، هرچقدر شباهت عامل مشغول شده به عامل مورد بررسی کمتر باشد، ارزش عکس‌العمل او بیشتر خواهد بود. در خلال توضیح مؤلفهٔ بعدی که تنوع اجتماعی است، با نحوه‌ی محاسبه‌ی شباهت دو عامل آشنا خواهیم شد.

تنوع اجتماعی[1] عامل‌های که به یک عامل عکس‌العمل نشان داده‌اند، به‌عنوان یک شاخص در میزان نفوذ آن عامل به حساب می‌آید. بدین صورت که به طور مثال هر چقدر تنوع اجتماعی ری‌توییت‌کنندگان یک عامل بیشتر باشد، جامعهٔ مخاطب گسترده‌تری در معرض پیام آن عامل قرار می‌گیرند. تنوع اجتماعی مجموعه‌ای از عامل‌ها با توجه به میزان شباهت آن‌ها با یکدیگر سنجیده می‌شود؛ بدین نحو که عامل‌ها مشابه به هم، در یک خوشه قرار می‌گیرند. تنوع اجتماعی ری‌توییت‌کنندگان یک عامل، زمانی بیشتر است که ری‌توییت‌کنندگان از نظر شباهت در تعداد خوشه‌های بیشتری قرار بگیرند. شباهت دو عامل مبتنی بر روابط دنبال‌کنندگی و ری‌توییت آن‌ها از یکدیگر محاسبه می‌شود.

$$f_{common}(a.b) = \text{تعداد دنبال شدگان مشترک} \quad (۲۳)$$

$$R_{common}(a.b) = \text{تعداد ری‌توییت شدگان مشترک} \quad (۲۴)$$

$$f_{direct}(a.b) = \begin{cases} 0 & (a\ don't\ follow\ b)\ AND\ (b\ don't\ follow\ a) \\ 1 & (a\ follows\ b)\ OR\ (b\ follows\ a) \\ 2 & (a\ follows\ b)\ AND\ (b\ follows\ a) \end{cases} \quad (۲۵)$$

$$R_{direct}(a.b) = \begin{cases} 0 & (a\ don't\ retweet\ b)\ AND\ (b\ don't\ retweet\ a) \\ 1 & (a\ retweets\ b)\ OR\ (b\ retweets\ a) \\ 2 & (a\ retweets\ b)\ AND\ (b\ retweets\ a) \end{cases} \quad (۲۶)$$

$$similarity(a,b) = (\alpha f_{common}(a,b) + \beta R_{common}(a,b) + \gamma f_{direct}(a,b) + \lambda R_{direct}(a,b)) \quad (۲۷)$$

همان‌طور که مشخص است، شباهت بین دو عامل بر اساس چهار ویژگی سنجیده می‌شود:

۱. تعداد عامل‌های که هر دو عامل آن‌ها را دنبال می‌کنند

۲. تعداد عامل‌های که هر دو عامل آن‌ها را ری‌توییت کرده‌اند.

۳. دنبال‌کردن یک عامل توسط عامل دیگر

۴. ری‌توییت کردن عامل توسط عامل دیگر

براین‌اساس اگر هیچ یک از دو عامل، یکدیگر را دنبال نکرده باشند، مقدار $f_{direct}(a.b)$ برابر صفر خواهد بود. اگر یکی از دو عامل، عامل دیگر را دنبال کرده باشد، این مقدار برابر با ۱ و در صورتیکه هر دو عامل یکدیگر را دنبال کرده باشند، این مقدار برابر ۲ خواهد بود. همچنین در صورتیکه هیچ یک از دو عامل یکدیگر را ری‌توییت نکرده باشند، مقدار $R_{direct}(a.b)$ برابر با صفر، در صورتیکه یکی از دو عامل، عامل دیگر را ری‌توییت کرده باشد، این مقدار برابر با ۱ و در صورتیکه ری‌توییت دو طرفه انجام شده باشد، این مقدار برابر ۲ خواهد بود.

با توجه به این فرمول، درصورتی‌که هیچ شباهتی بین عامل تولیدکننده پیام و عاملی که به آن عکس‌العمل نشان داده وجود نداشته باشد، امتیاز پیام بی‌نهایت خواهد شد. لذا جهت جلوگیری از این مشکل، برای این دسته از عامل‌ها، به‌جای صفر قرار دادن $similarity$، مقدار آن را برابر با $\min(similarity)/2$ قرار می‌دهیم. در نهایت با نرمالیزه کردن مقادیر این چهار متغییر بین صفر و یک و یافتن ضرایب α، β، γ و λ مبتنی بر آنتروپی مقادیر هر یک از چهار متغییر، مقدار $similarity(a,b)$ محاسبه می‌شود.

حال با به‌دست‌آمدن شباهت بین هر دو عامل، ماتریس

[1] Social Diversity (SD)

شباهت کل عامل‌ها به وجود آمده است. لذا با اعمال الگوریتم خوشه‌بندی طیفی بر روی ماتریس شباهت، خوشه‌بندی عامل‌ها و تعداد خوشه‌های آن‌ها مشخص می‌شود و تنوع اجتماعی به‌صورت زیر به دست خواهد آمد:

$SD(u) =$

(۲۸) تعداد خوشه های موجود در ماتریس شباهت عامل‌ها مشغول

۵. ارزیابی مدل ارائه شده

۵-۱. الگوریتم‌های مشابه

جهت ارزیابی روش ارائه شده، ما به مقایسه کارایی آن با ۴ الگوریتم پرداخته‌ایم. اولین الگوریتم ($TwitterFL$) [43] جهت رتبه‌بندی عامل، مجموع دنبال‌کنندگان عامل با دنبال‌کنندگان کل ری‌توییت‌کنندگان او را مورد توجه قرار می‌دهد. در الگوریتم دوم ($InfluenceFinder$)[44]، جمع وزنی تعداد دنبال‌کنندگان، تعداد ریتوییت‌های دریافتی و تعداد لایک‌های دریافتی عامل به صورتی که وزن هر یک به ترتیب برابر با ۰/۵، ۰/۳۳ و ۰/۱۷ باشد، مورد استفاده قرار می‌گیرد. الگوریتم سوم (SNP)[19] نیز با معرفی دو معیار نسبت تعامل[1] و نسبت ری‌توییت و منشن[2] به محاسبه‌ی معیاری به نام پتانسیل شبکه‌سازی اجتماعی[3] می‌پردازد.

$$Ir(i) = \frac{(\text{تعداد منشن کنندگان})+(\text{تعداد ری‌توییت کنندگان})}{\text{تعداد دنبال کنندگان}} \quad (۲۹)$$

$$RMr(i) = \frac{\begin{pmatrix}(\text{توییت های پاسخ داده شده عامل})\\+\\(\text{تعداد توییت های ری‌توییت شده عامل})\end{pmatrix}}{\text{تعداد توییت های عامل}} \quad (۳۰)$$

$$SNP(i) = \frac{Ir(i)+RMr(i)}{2} \quad (۳۱)$$

روش ارائه شده با عنوان $Twitterrank$ [29]به‌عنوان یکی از روش‌های مرجع که در بسیاری از مقاله‌ها مورد مقایسه قرار گرفته است، به‌عنوان الگوریتم چهارم مورد استفاده قرار گرفته است. این الگوریتم مبتنی بر رتبه صفحه کار می‌کند با این تفاوت که جهت یافتن شباهت عامل‌ها، محتوای آن‌ها را هم در نظر می‌گیرد.

۵-۲. معیارهای مقایسه

جهت مقایسه الگوریتم ارائه شده با الگوریتم‌هایی که در زیربخش قبلی به آن‌ها اشاره شد، نیاز است تا معیارهایی مورد استفاده قرار گیرند که به ارزیابی رتبه‌بندی عامل‌ها می‌پردازد. با توجه به اینکه در این مسئله ما رتبه‌بندی صحیح عامل‌ها را نداریم، لذا نمی‌توانیم از معیارهایی که مبتنی بر مقایسه‌ی رتبه‌بندی حاصل از اعمال هر الگوریتم با رتبه‌بندی صحیح هستند (مانند precision، recall، F-measure، accuracy و ...)، استفاده کنیم. ازاین‌رو باید به دنبال معیارهایی باشیم که به‌صورت بدون ناظر و یا به‌صورت اکتشافی به ارزیابی رتبه‌بندی عامل‌ها می‌پردازند. معیارهای زیر، معیارهایی هستند که این ویژگی را دارند:

۱. درصد امتیاز یکتا[4][45]: یک روش امتیاز دهی مناسب، روشی‌است که امکان ایجاد تمایز بین نفوذ عامل‌ها را تا بالاترین حد ممکن داشته باشد. به بیان دیگر، داشتن مقادیر تکراری نفوذ برای عامل‌ها، به معنای کارایی کمتر الگوریتم است. فرض کنید ξ تعداد عامل‌ها با میزان نفوذ یکتا است. درصد امتیاز یکتا η به‌صورت زیر محاسبه می‌شود:

$$\eta = \frac{\xi}{N} * 100\% \quad (۳۲)$$

۲. انتشار اطلاعات[5] [۴۷]: این معیار به سنجش تعداد عامل‌های که با انتخاب K عامل اولیه، مورد نفوذ واقع شده‌اند، می‌پردازد. انتشار اطلاعات بالاتر نمایانگر کارایی بالاتر الگوریتم است. جهت مقایسه کارایی الگوریتم‌ها با یکدیگر، ۱۰۰ عامل برتر با توجه به هر الگوریتم انتخاب می‌شود و تعداد عامل‌ها مورد نفوذ واقع شده، از طریق آن ۱۰۰ نفر، با بهره‌گیری از الگوریتم آبشار مستقل[6] سنجش می‌شود.

۳. طول درخت نفوذ[7] : پس از احصاء عامل‌های که از طریق ۱۰۰ عامل برتر، مورد نفوذ واقع شده‌اند، می‌توان با سنجش مجموع فاصلهٔ آن‌ها با ۱۰۰ عامل برتر، طول درخت نفوذ را محاسبه نمود. کمتر بودن طول درخت نفوذ، بیانگر کارایی بیشتر الگوریتم است.

۴. بهینگی نفوذ: با توجه به تعداد عامل‌ها مورد نفوذ واقع شده و طول درخت نفوذ، می‌توان بهینگی نفوذ را به‌صورت زیر تعریف کرد:

$$\text{بهینگی نفوذ} = \frac{\text{انتشار اطلاعات}}{\text{طول درخت نفوذ}} \quad (۳۳)$$

۵. پوشش نفوذ کیفی: با معیار قرار دادن الگوریتم ارائه شده،

[1] Intractor Ratio (Ir)
[2] Retweet and Mention Ratio (RMr)
[3] Social Networking Potential (SNP)
[4] Duplication percentage
[5] Information Spread
[6] Independent Cascade
[7] Length of Influence Tree

نسبت بین مجموع امتیاز عامل‌ها مورد نفوذ واقع شده به مجموع امتیاز کل عامل‌ها، معیاری است از اینکه هر الگوریتم چه میزان از پوشش نفوذ را داراست. به بیان دیگر، جهت سنجش این معیار به‌صورت زیر عمل می‌کنیم:

- مشخص‌کردن لیست عامل‌های که اطلاعات به دست آن‌ها رسیده بر اساس هر یک از الگوریتم‌ها
- محاسبه مجموع امتیاز نفوذ عامل‌ها لیست بالا بر اساس الگوریتم ارائه شده
- محاسبه‌ی پوشش نفوذ کیفی بر اساس فرمول زیر:

$$\mu = \frac{\text{مجموع امتیازات عامل‌هایی که اطلاعات به دست آن ها رسیده}}{\text{مجموع امتیازات کل عامل‌ها}} \quad (۳۴)$$

۳-۵. مجموعه دادگان

در این تحقیق جهت ارزیابی مدل مفهومی و محاسباتی ارائه شده، از داده‌های واقعی استفاده شده است. با توجه به این‌که از بین مجموعه دادگان استاندارد موجود که در مقالات مورد استفاده قرار می‌گیرند، دادگانی وجود نداشت که ویژگی‌های مورداستفادهٔ روش ارائه شده را پوشش دهد و یا حداقل، این ویژگی‌ها را بتوان از آن دادگان استخراج کرد، لذا ما به تولید مجموعه دادگان مرتبط با سه رخداد که حول آن‌ها در شبکه‌های اجتماعی در سال ۱۴۰۰ جریانی شکل گرفته است، اقدام نمودیم. سه رخداد قطع برق، عید قربان و تعویق کنکور کارشناسی ارشد، رخدادهایی بودند که داده‌های آن‌ها خزش و تحلیل شدند. ویژگی‌های این مجموعه‌های دادگان در جدول ۳ زیر بیان شده است.

جدول ۳: ویژگی‌های مجموعه دادگان ساخته شده

عنوان	بازهٔ زمانی	تعداد عامل‌ها فعال	تعداد عامل‌ها قابل خزش	تعداد یال‌ها در گراف دنبال‌کنندگی	میانگین درجه نودها	تعداد توییت	تعداد ری‌توییت	تعداد پاسخ	تعداد نقل قول
تعویق کنکور کارشناسی ارشد	۱۵ تیر تا ۱۰ مرداد	۱۵۵۳	۱۳۵۱	۳۰۸۱۲	۲۳/۴۸	۲۶۶۵۵	۴۶۷۸۳۸	۴۹۹۳	۲۰۰۵
عید قربان	۱۸ تیر تا ۱ مرداد	۲۵۲۳۸	۲۳۷۱۵	۷۷۲۶۷۴۴	۳۲۹/۴۱	۱۵۷۴۳	۵۲۵۶۹	۳۲۴۹	۴۰۷
قطع برق	۱ خرداد تا ۱ شهریور	۳۳۱۹۲	۳۱۲۶۸	۹۴۵۴۳۸۵	۳۰۳/۷۲	۲۵۰۸۵	۱۲۵۰۵۲	۲۶۹۹	۱۷۰۱

با توجه به بازهٔ زمانی هر یک از مجموعه‌های دادگان و تغییرات رخ‌داده شده در حساب‌های عاملی عامل‌ها (حذف شدن، تعلیق شدن و تبدیل به خصوصی شدن)، در زمان خزش اطلاعات عامل‌ها فعال در هر موضوع، اطلاعات برخی عامل‌ها قابل خزش نبود. لذا در جدول ۳، تعداد عامل‌ها قابل خزش نیز مشخص شده است. فرایند ایجاد مجموعه دادگان در هر موضوع، شامل گام‌های زیر است:

۱. مشخص‌کردن کلیدواژه‌های مرتبط با هر موضوع
۲. خزش محتوای حاوی کلیدواژه‌های مشخص شده
۳. مشخص‌کردن لیست عامل‌ها فعال در موضوع (عامل‌های که به ایجاد توییت، ری‌توییت، پاسخ و نقل قول پرداخته‌اند.)
۴. خزش اطلاعات عاملی عامل‌ها فعال (شامل تعداد توییت، تعداد دنبال‌کنندگان، تعداد دنبال‌شوندگان)
۵. خزش لیست دنبال‌شوندگان عامل‌ها فعال
۶. تشکیل گراف دنبال‌کنندگی بین عامل‌ها فعال (عامل‌های که در موضوع مورد نظر فعالیتی نداشته‌اند و در لیست دنبال‌شوندگان عامل‌ها فعال قرار دارند، در گراف دنبال‌کنندگی نخواهند آمد.) در این گراف یک یال جهت‌دار از هر دنبال‌کننده به هر دنبال‌شونده وجود دارد.
۷. ایجاد لیست عامل‌های ری‌توییت‌کنندهٔ یک توییت
۸. ایجاد لیست عامل‌های پاسخ‌دهنده به یک توییت
۹. ایجاد لیست عامل‌های نقل قول کنندهٔ یک توییت

۴-۵. نتایج ارزیابی

۱-۴-۵. مجموعه دادگان تعویق کنکور ارشد

شکل ۴: گراف‌های مرتبط با مجموعه دادگان تعویق کنکور ارشد. سمت راست: گراف ریتوییت - سمت چپ: گراف دنبال‌کنندگی

این مجموعه دادگان متعلق به کمپینی است که برخی عامل‌ها با هدف تعویق کنکور کارشناسی ارشد به دلیل وجود ویروس کرونا راه‌اندازی کردند. گراف دنبال‌کنندگی و ری‌توییت عامل‌ها (شکل ۴) در این کمپین به‌صورت زیر است. در گراف ری‌توییت نیز یک یال وزن‌دار بین ری‌توییت‌کننده و ری‌توییت‌شونده وجود دارد که وزن آن، نشانگر تعداد ریتوییت کردن می‌باشد.

جدول ۴: مقایسه عملکرد الگوریتم‌ها روی مجموعه دادگان تعویق کنکور ارشد

ردیف	عنوان الگوریتم	درصد نفوذ یکتا	انتشار اطلاعات	طول درخت نفوذ	بهینگی نفوذ ($*10^{-2}$)	پوشش نفوذ کیفی
۱	الگوریتم ما	**۹۹/۲۹**	۹۲۰	۱۴۸۶۵۲	۰/۶۱	**۰/۶۵۵**
۲	$TwitterFL$	۵۱/۳۲	۹۲۳	**۱۴۶۲۲۷**	**۰/۶۳**	۰/۶۴۲
۳	SNP	۱۲/۸۱	۹۵۲	۲۵۴۵۷۷	۰/۳۷	۰/۵۹۱
۴	$InfluenceFinder$	۳۷/۶۶	**۹۶۵**	۲۱۰۰۲۴	۰/۴۵	۰/۶۲۴
۵	$twitterRank$	۷۴/۷۵	۹۰۸	۲۱۳۲۵۲	۰/۴۲	۰/۵۸۰

در گراف سمت راست، ساختار ارتباطی عامل‌ها مبتنی بر روابط دنبال‌کنندگی به نمایش گذاشته شده است. همان‌طور که مشخص است، عامل‌ها با اعمال الگوریتم ماژولاریتی، به عنوان معروفترین الگوریتم خوشه‌بندی گراف، به ۵ دسته تقسیم شده‌اند که با رنگ‌های مختلف نمایش داده شده‌اند. با نگاهی به گراف سمت چپ که گراف ری‌توییت‌کنندگاناست، پی به این می‌بریم که به دلیل ری‌توییت شدن عامل‌های یک انجمن خاص توسط دیگر عامل‌ها، در گراف ری‌توییت، دیگر انجمن‌ها با این انجمن خاص ترکیب شده‌اند. در این گراف، عامل‌هایی که با رنگ مشکی پررنگ نمایش داده شده‌اند، عامل‌هایی هستند که امکان گرفتن اطلاعات حساب عاملی آن‌ها وجود نداشته است و در گراف دنبال‌کنندگی وجود ندارند. در جدول ۴ نتایج اعمال الگوریتم‌های مختلف روی این مجموعه دادگان نشان‌داده‌شده است.

۲-۴-۵. مجموعه دادگان قطع برق

این مجموعه دادگان مربوط به بازهٔ زمانی‌ای است که به دلیل افزایش شدید دما و همچنین فعالیت ماینرهای رمزارز، در کشور با قطعی برق روبرو بودیم. گراف ری‌توییت و دنبال‌کنندگی عامل‌ها در این کمپین در شکل ۵ مشخص شده است. در جدول ۵ نتایج الگوریتم‌های مختلف روی این مجموعه دادگان

نشان‌داده‌شده است.

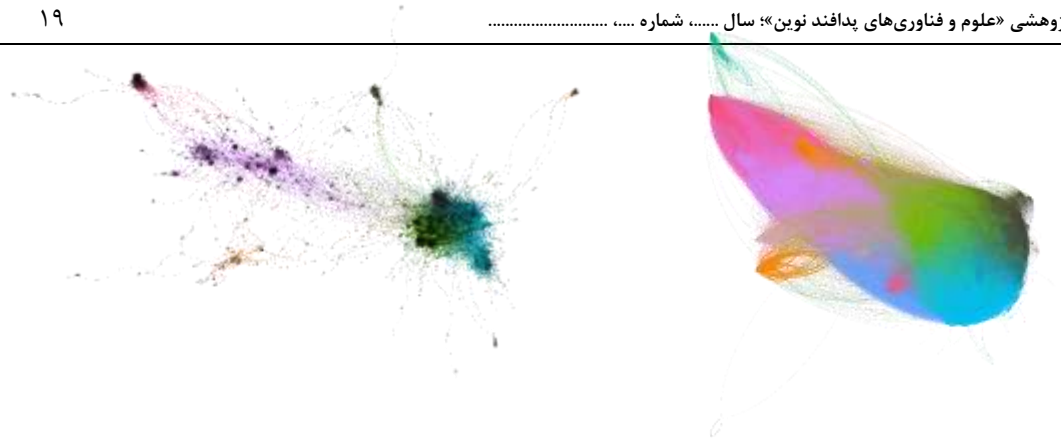
شکل ۵: گراف‌های مرتبط با مجموعه دادگان قطع برق. سمت راست: گراف ریتوییت - سمت چپ: گراف دنبال‌کنندگی

جدول ۵: مقایسه عملکرد الگوریتم‌ها روی مجموعه دادگان قطع برق

ردیف	عنوان الگوریتم	درصد نفوذ یکتا	انتشار اطلاعات	طول درخت نفوذ	بهینگی نفوذ ($*10^{-2}$)	پوشش نفوذ کیفی
۱	الگوریتم ما	**۹۵/۴۹**	۲۹۵۳۱	۶۸۸۰۵۷۲	۰/۴۲	۰/۹۴۸۴
۲	$TwitterFL$	۱۲/۸۰	۲۹۴۴۱	۶۸۲۷۹۷۹	۰/۴۳	۰/۹۴۰۰
۳	SNP	۴/۴۳	**۲۹۶۰۰**	۶۱۲۷۷۱۱	۰/۴۸	**۰/۹۵۰۵**
۴	$InfluenceFinder$	۱۱/۶۳	۲۹۵۷۵	۶۶۷۱۲۸۱	۰/۴۴	۰/۹۴۷۰
۵	$twitterRank$	۴۰/۷۰	۲۹۵۱۴	**۶۰۴۲۰۰**	**۴/۸۸**	۰/۹۴۸۰

۳-۴-۵. مجموعه دادگان عید قربان

این مجموعه دادگان نیز مربوط به بازهٔ زمانی حوالی عید سعید قربان است. لازم به ذکر است که در خزش اطلاعات مربوط به این موضوع، حجم قابل‌توجهی از عامل‌ها، عامل‌ها عرب‌زبان بوده‌اند. گراف ری‌توییت و دنبال‌کنندگی عامل‌ها در این کمپین به‌صورت زیر (شکل ۶) است. در جدول ۶ نتایج الگوریتم‌های مختلف روی این مجموعه دادگان نشان‌داده‌شده است.



جدول ۶: مقایسه عملکرد الگوریتم‌ها روی مجموعه دادگان عید قربان

شکل ۶: گراف‌های مرتبط با مجموعه دادگان عید قربان. سمت راست: گراف ریتوییت - سمت چپ: گراف دنبال‌کنندگی

ردیف	عنوان الگوریتم	درصد نفوذ یکتا	انتشار اطلاعات	طول درخت نفوذ	بهینگی نفوذ (*10^{-2})	پوشش نفوذ کیفی
۱	الگوریتم ما	**۹۷/۲۳**	۲۱۵۹۲	۵۲۶۴۳۴۱	۰/۴۱	**۰/۹۴۰۱**
۲	$TwitterFL$	۱۴/۸۷	۲۱۵۴۸	**۴۲۹۳۹۱۳**	**۰/۵۰**	۰/۹۳۷۰
۳	SNP	۵/۳۰	۲۱۵۷۹	۴۳۴۹۹۴۷	۰/۴۹	۰/۹۳۷۹
۴	$InfluenceFinder$	۱۵/۲۰	**۲۱۶۳۹**	۵۰۱۴۹۷۰	۰/۴۳	۰/۹۳۸۱
۵	$twitterRank$	۹/۲۳	۲۱۶۰۵	۴۶۷۳۱۱۶	۰/۴۶	۰/۹۳۷۱

۶. نتیجه‌گیری

مطالعات در حوزه ارزیابی کاربران شبکه‌های اجتماعی برخط با اهداف مختلفی صورت گرفته است. بعضی از این کاربران، به‌عنوان عامل عملیات نفوذ شناختی ایفای نقش می‌کنند. ارزیابی این کاربران عامل، هدف این تحقیق بوده است. بر اساس نقشی که یک عامل می‌تواند در موفقیت یک عملیات نفوذ شناختی داشته باشد، مؤلفه‌ها و شاخص‌هایی در نظر گرفته شد. با توجه به شاخص‌های مورد استفاده در این مدل، مجموعه دادگانی که شامل همهٔ این شاخص‌ها باشد، وجود ندارد و ما سه مجموعه دادگان حاوی شاخص‌های مورد نظر تولید کردیم. نتایج به‌دست‌آمده نشانگر این مطلب است که مدل ارائه شده در کنار قابلیت تفسیرپذیری و عدم نیاز به داده‌های آموزشی، دارای عملکردی قابل مقایسه با روش‌های قبلی است.

مدل ارائه شده بر اساس مدل مفهومی ارزیابی توان شبکه‌ای پیام یا کاربر در شبکه‌ی اجتماعی مبتنی بر نظریه‌های ارتباطاتی شکل گرفته است. لذا خروجی رتبه‌بندی این مدل چه برای پیام و چه برای کاربر، معنادار و تفسیرپذیر است. در عین حال آزمایش‌های انجام شده ثابت می‌کند که این روش در مقایسه با دیگر روش‌ها، از لحاظ معیارهای ارزیابی الگوریتم نتایج بهتر و قابل مقایسه‌ای دارد.

۷. مراجع

[1] P. D. Allen, *Information Operations Planning*. Artech House, 2007.

[2] W. Lu, “Computational social influence : models, algorithms, and applications,” University of British Columbia, 2016.

[3] J. Stubbs and C. Bing, “Exclusive: Iran-based political influence operation - bigger, persistent, global,” 2018. https://www.reuters.com/article/us-usa-iran-facebook-exclusive/exclusive-iran-based-political-influence-operation-bigger-persistent-global-idUSKCN1LD2R9 (accessed Nov. 01, 2018).

[4] J. Publication, “Information Operations Joint Publication 3-13,” no. November 2012.

[5] L. Ben Jabeur, L. Tamine, and M. Boughanem, “Active microbloggers: Identifying influencers, leaders and discussers in microblogging networks,” in *Lecture Notes in Computer Science*

2013, doi: 10.1016/j.ipm.2013.06.003.

[18] J. Zhang, R. Zhang, J. Sun, Y. Zhang, and C. Zhang, "TrueTop: A Sybil-Resilient System for User Influence Measurement on Twitter," *IEEE/ACM Trans. Netw.*, vol. 24, no. 5, 2016, doi: 10.1109/TNET.2015.2494059.

[19] I. Anger and C. Kittl, "Measuring influence on Twitter," 2011, doi: 10.1145/2024288.2024326.

[20] T. Noro, F. Ru, F. Xiao, and T. Tokuda, "Twitter user rank using keyword search," *Front. Artif. Intell. Appl.*, vol. 251, 2013, doi: 10.3233/978-1-61499-177-9-31.

[21] M. S. Srinivasan, S. Srinivasa, and S. Thulasidasan, "Exploring celebrity dynamics on twitter," 2013, doi: 10.1145/2528228.2528242.

[22] M. S. Srinivasan, S. Srinivasa, and S. Thulasidasan, "A comparative study of two models for celebrity identification on twitter," *Proc. 20th Int. Conf. Manag. Data*, 2014.

[23] S. Kong and L. Feng, "A tweet-centric approach for topic-specific author ranking in micro-blog," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2011, vol. 7120 LNAI, no. PART 1, doi: 10.1007/978-3-642-25853-4_11.

[24] D. Hatcher, G. S. Bawa, and B. de Ville, "How you can identify influencers in SAS R Social media analysis (and why it matters)," in *SAS Global Forum*, 2011, pp. 4–7.

[25] Z. Y. Ding, Y. Jia, B. Zhou, Y. Han, L. He, and J. F. Zhang, "Measuring the spreadability of users in microblogs," *J. Zhejiang Univ. Sci. C*, vol. 14, no. 9, 2013, doi: 10.1631/jzus.CIIP1302.

[26] A. Silva, S. Guimarães, W. Meira, and M. Zaki, "ProfileRank: Finding relevant content and influential users based on information diffusion," 2013, doi: 10.1145/2501025.2501033.

[27] P. Y. Huang, H. Y. Liu, C. T. Lin, and P. J. Cheng, "A diversity-dependent measure for discovering influencers in social networks," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2013, vol. 8281 LNCS, doi: 10.1007/978-3-642-45068-6_32.

[28] D. Liu, Q. Wu, and W. Han, "Measuring micro-blogging user influence based on user-tweet interaction model," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2013, vol. 7929 LNCS, no. PART 2, doi: 10.1007/978-3-642-38715-9_18.

[29] J. Weng, E. P. Lim, J. Jiang, and Q. He, "Twitterrank: Finding topic-sensitive influential twitterers," *Proc. 3rd ACM Int. Conf. Web Search Data Min. (WSDM 2010)*, pp. 261–270, 2010, doi: 10.1145/1718487.1718520.

[30] A. Aleahmad, P. Karisani, M. Rahgozar, and F. Oroumchian, "OLFinder: Finding opinion leaders *(including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2012, vol. 7608 LNCS, doi: 10.1007/978-3-642-34109-0_12.

[6] B. Krishnamurthy, P. Gill, and M. Arlitt, "A few chirps about Twitter," 2008, doi: 10.1145/1397735.1397741.

[7] J. J. F. Forest, *Influence Warfare: How Terrorists and Governments Fight to Shape Perceptions in a War of Ideas: How Terrorists and Governments Fight to Shape Perceptions in a War of Ideas*. ABC-CLIO, 2009.

[8] L. Page, S. Brin, R. Motwani, and T. Winograd, "The PageRank Citation Ranking: Bringing Order to the Web," *World Wide Web Internet Web Inf. Syst.*, vol. 54, no. 1999–66, 1998, doi: 10.1.1.31.1768.

[9] Y. Yamaguchi, T. Takahashi, T. Amagasa, and H. Kitagawa, "TURank: Twitter user ranking based on user-tweet graph analysis," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2010, vol. 6488 LNCS, doi: 10.1007/978-3-642-17616-6_22.

[10] R. Nagmoti, A. Teredesai, and M. De Cock, "Ranking approaches for microblog search," in *Proceedings - 2010 IEEE/WIC/ACM International Conference on Web Intelligence, WI 2010*, 2010, vol. 1, doi: 10.1109/WI-IAT.2010.170.

[11] T. Majer and M. Šimko, "Leveraging microblogs for resource ranking," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2012, vol. 7147 LNCS, doi: 10.1007/978-3-642-27660-6_42.

[12] B. Hajian and T. White, "Modelling influence in a social network: Metrics and evaluation," 2011, doi: 10.1109/PASSAT/SocialCom.2011.118.

[13] A. Khrabrov and G. Cybenko, "Discovering influence in communication networks using dynamic graph analysis," 2010, doi: 10.1109/SocialCom.2010.48.

[14] Z. Ding, Y. Jia, B. Zhou, and Y. Han, "Mining topical influencers based on the multi-relational network in micro-blogging sites," *China Commun.*, vol. 10, no. 1, 2013, doi: 10.1109/CC.2013.6457533.

[15] D. M. Romero, W. Galuba, S. Asur, and B. A. Huberman, "Influence and passivity in social media," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2011, vol. 6913 LNAI, no. PART 3, pp. 18–33, doi: 10.1007/978-3-642-23808-6_2.

[16] Z. Yin and Y. Zhang, "Measuring pair-wise social influence in microblog," 2012, doi: 10.1109/SocialCom-PASSAT.2012.10.

[17] D. Gayo-Avello, "Nepotistic relationships in Twitter and their impact on rank prestige algorithms," *Inf. Process. Manag.*, vol. 49, no. 6,

in online social networks," *J. Inf. Sci.*, vol. 42, no. 5, 2016, doi: 10.1177/0165551515605217.

[31] A. Pal and S. Counts, "Identifying topical authorities in microblogs," 2011, doi: 10.1145/1935826.1935843.

[32] M. Montangero and M. Furini, "TRank: Ranking Twitter users according to specific topics," 2015, doi: 10.1109/CCNC.2015.7158074.

[33] X. Li, S. Cheng, W. Chen, and F. Jiang, "Novel user influence measurement based on user interaction in microblog," 2013, doi: 10.1145/2492517.2492635.

[34] U. Ishfaq, H. U. Khan, S. Iqbal, and M. Alghobiri, "Finding influential users in microblogs: state-of-the-art methods and open research challenges," *Behav. Inf. Technol.*, 2021, doi: 10.1080/0144929X.2021.1915384.

[35] L. Qi, Y. Huang, L. Li, and G. Xu, "Learning to rank domain experts in microblogging by combining text and non-text features," 2015, doi: 10.1109/BESC.2015.7365953.

[36] M. Yu, W. Yang, W. Wang, and G. W. Shen, "Information influence measurement based on user quality and information attribute in microblogging," 2016, doi: 10.1109/ICCSN.2016.7586594.

[37] G. S. Mahalakshmi, K. Koquilamballe, and S. Sendhilkumar, "Influential detection in twitter using tweet quality analysis," 2017, doi: 10.1109/ICRTCCM.2017.62.

[38] X. Luo, L. Zhang, Y. Yi, R. Xue, and D. Jiang, "The key user discovery model based on user importance calculation," *Int. J. Comput. Sci. Eng.*, vol. 21, no. 2, 2020, doi: 10.1504/ijcse.2020.10027436.

[39] C. Lee, H. Kwak, H. Park, and S. Moon, "Finding influentials based on the temporal order of information adoption in Twitter," 2010, doi: 10.1145/1772690.1772842.

[40] C. Sun, L. Zhang, and Q. Li, "Who are influentials on micro-blogging services: Evidence from social network analysis," 2013.

[41] J. Yuan, L. Li, L. Luo, and M. Huang, "Topology-based algorithm for users' influence on specific topics in micro-blog," *J. Inf. Comput. Sci.*, vol. 10, no. 8, 2013, doi: 10.12733/jics20102229.

[42] and Y. J. Z. Y. B. Zhuang, Z. H. Li, "Identification of influencers in online social networks: measuring influence considering multidimensional factors exploration," *HELIYON*, vol. 7, no. 4, 2021.

[43] K. K. Darsipudi, "Influential User Detection," 2017. https://github.com/krishnakartik1/influentialUser Detection.

[44] "Influence Finder," 2017. https://github.com/sapansanu/InfluenceFinder.

[45] and Z. X. G. Wang, W. Jiang, J. Wu, "Fine-grained feature-based social influence evaluation in online social networks," *IEEE Trans. Parallel Distrib. Syst.*, vol. 25, no. 9, pp. 2286–2296, 2014.